



REFERENCE ONLY

UNIVERSITY OF LONDON THESIS

Degree

PhD

Year

2005

Name of Author

FLETCHER, C. G.

COPYRIGHT

This is a thesis accepted for a Higher Degree of the University of London. It is an unpublished typescript and the copyright is held by the author. All persons consulting the thesis must read and abide by the Copyright Declaration below.

COPYRIGHT DECLARATION

I recognise that the copyright of the above-described thesis rests with the author and that no quotation from it or information derived from it may be published without the prior written consent of the author.

LOANS

Theses may not be lent to individuals, but the Senate House Library may lend a copy to approved libraries within the United Kingdom, for consultation solely on the premises of those libraries. Application should be made to: Inter-Library Loans, Senate House Library, Senate House, Malet Street, London WC1E 7HU.

REPRODUCTION

University of London theses may not be reproduced without explicit written permission from the Senate House Library. Enquiries should be addressed to the Theses Section of the Library. Regulations concerning reproduction vary according to the date of acceptance of the thesis and are listed below as guidelines.

- A. Before 1962. Permission granted only upon the prior written consent of the author. (The Senate House Library will provide addresses where possible).
- B. 1962 - 1974. In many cases the author has agreed to permit copying upon completion of a Copyright Declaration.
- C. 1975 - 1988. Most theses may be copied upon completion of a Copyright Declaration.
- D. 1989 onwards. Most theses may be copied.

This thesis comes within category D.



This copy has been deposited in the Library of UCL



This copy has been deposited in the Senate House Library, Senate House, Malet Street, London WC1E 7HU.

Investigations into Seasonal Predictability of North Atlantic Winter Climate

Christopher G. Fletcher

*Thesis submitted for the degree
of Doctor of Philosophy of the
University of London*

Mullard Space Science Laboratory
Department of Space and Climate Physics
University College London
April 2005

UMI Number: U592015

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U592015

Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

For my family.

Abstract

Skilful seasonal predictability of European climate would bring widespread socio-economic benefits. However, little useful skill has been identified to date. This work extends prior research to show evidence for significant skill in predicting the winter North Atlantic Oscillation (NAO). The research divides into two topics. First, the work clarifies what is the best lagged predictor for the winter NAO by standardising the assessment of four previously published lagged NAO predictors. A new NAO predictor – the zonal gradient in summer northern hemisphere (NH) subpolar air temperature – is examined. This predictor outperforms other NAO predictors over assessment periods out to 100 years. This finding suggests that it is NH subpolar regions rather than the midlatitudes or the tropics which provide the best NAO lagged predictability. A physical mechanism linking summer NH subpolar climate and the winter NAO is proposed. Summertime subpolar atmospheric perturbations lead a pattern of North Atlantic sea surface temperature which persists into autumn. This pattern could feed back onto the atmosphere to influence the sign and magnitude of the winter NAO. Second, the study explores further the mechanisms which underpin the observed NAO predictability – in particular how summer NH snow cover links to the winter NAO – by using a coupled general circulation model (CGCM). A validation of the CGCM snow cover representation 1972-2002 is presented. The CGCM captures well the observed annual cycle and spatial distribution of NH snow cover. However, the CGCM exhibits critical deficiencies in the seasonal interaction between snow cover and the atmosphere. The CGCM exhibits a significant link between summer snow cover and the winter NAO 1972-2002. However, this link is nonstationary during the twentieth century and does not function through the mechanism seen in observations. Winter NAO predictability in a single CGCM integration cannot therefore be distinguished from internal model variability.

Contents

List of Tables	8
List of Figures	10
Acknowledgements	18
1 Introductory Material	19
1.1 Introduction	19
1.2 Seasonal climate forecasting	20
1.2.1 Physical basis	20
1.2.2 Dynamical SCF	25
1.2.3 Empirical SCF	27
1.2.4 Dynamical versus empirical SCF	29
1.3 Potential benefits of SCF	31
1.3.1 Benefits of seasonal forecasting for the winter NAO	31
1.3.2 User groups	33
1.4 Seasonal predictability of the winter North Atlantic Oscillation	35
1.4.1 Definition and background	35
1.4.2 The winter season	36
1.4.3 NAO or AO?	38
1.4.4 Boundary forcing of the NAO_{DJF}	40
1.4.5 Lagged NAO_{DJF} predictability and skill	42
1.4.6 Research questions	45
1.5 GCM representation of NAO links to prior Northern Hemisphere snow cover	45
1.5.1 The relationships between snow cover and climate	46
1.5.2 GCM simulation of snow cover	51
1.5.3 GCM simulation of snow/atmosphere links	53
1.5.4 Research questions	55
1.6 Summary	56

2	Data Selection	57
2.1	Introduction	57
2.2	MSLP data	58
2.3	NAO indices	62
2.4	Near-surface air temperature	64
2.5	Sea surface temperature	70
2.6	Snow cover	75
2.7	Summary	79
3	What is the Best Lagged Predictor of the Winter NAO?	80
3.1	Introduction	80
3.2	Methodology	81
3.2.1	Data	81
3.2.2	Trends	85
3.2.3	Hindcast methodology	87
3.3	NAO hindcast skill	90
3.3.1	1900-2001	90
3.3.2	1950-2001	92
3.3.3	1972-2001	93
3.3.4	Stationarity	93
3.3.5	Data set dependence	94
3.4	Summer ΔT influence on upcoming NAO_{DJF}	95
3.4.1	Relationship between NAO_{DJF} lagged predictors	95
3.4.2	Role of prior winter NAO_{DJF}	97
3.4.3	Role of lagged snow cover	98
3.4.4	Physical basis for summer ΔT influence on upcoming winter NAO _{DJF}	98
3.5	Discussion	105
3.6	Summary and conclusions	108

4	Coupled GCM Representation of Observed Snow Cover and Associated Atmospheric Variability 1972-2002	109
4.1	Introduction	109
4.2	ECHAM4/OPYC3	111
4.3	Methodology	114
4.3.1	Correlation	114
4.3.2	Low-frequency variability	114
4.4	Comparison of observed and ECHAM4 snow cover variability	116
4.4.1	Derivation of ECHAM4 snow cover indices	116
4.4.2	Spatial variability	118
4.4.3	Temporal variability	120
4.5	Climatic response to <i>in situ</i> snow anomalies	128
4.5.1	Atmospheric temperature response	129
4.5.2	Surface flux response	132
4.6	Discussion	139
4.7	Summary and conclusions	140
5	Coupled GCM Representation of Observed NAO_{DJF} Predictability from NH Summer Snow Cover	142
5.1	Introduction	142
5.2	Methodology	143
5.2.1	Lead/lag correlation analysis	143
5.3	ECHAM4 NAO_{DJF} variability	144
5.4	ECHAM4 NAO_{DJF} predictability from lagged NH snow cover . . .	148
5.4.1	Lead/lag correlation analysis	148
5.4.2	NAO_{DJF} predictability in ECHAM4	150
5.5	Physical basis for NAO_{DJF} predictability in ECHAM4	153
5.5.1	Subpolar zonal temperature gradients	153
5.5.2	North Atlantic sector response	161
5.6	ECHAM4 stationarity 1904-2002	163
5.6.1	NAO_{DJF} predictability	164

5.6.2	Physical basis for NAO_{DJF} predictability	165
5.7	Discussion	168
5.8	Limitations of CGCMs for snow/atmosphere modelling	169
5.9	Summary and conclusions	171
6	Conclusions and Future Work	173
6.1	Introduction	173
6.2	Best lagged NAO_{DJF} predictor (Chapter 3)	173
6.3	Coupled GCM snow validation (Chapter 4)	176
6.4	NAO_{DJF} predictability from snow cover in a coupled GCM (Chapter 5)	179
6.5	Final remarks	182
	References	183

List of Tables

1.1	Summary comparison of empirical and dynamical methods used for SCF. After Lloyd-Hughes (2002).	29
2.1	Global analyses of sea level pressure data. Data set titles in brackets are those given by Smith and Reynolds (2004b) and allow comparison with Figure 2.1.	60
2.2	Global analyses of 2 m air temperature data.	68
2.3	Global analyses of sea surface temperature data.	74
2.4	Summary of the main data sets used in this study.	79
3.1	Summary of current NAO seasonal prediction studies examined in this Chapter.	84
3.2	Data sets used to compute NAO_{DJF} predictors in this study. NNR denotes NCEP/NCAR reanalysis data.	88
3.3	Skill values from cross-validated NAO hindcasts. Data set abbreviations: Had = HadISST, CRUT = CRUTEM2, Brn/Rut = Combined Brown and Rutgers snow cover index, NCEP = NCEP/NCAR Reanalysis, Rut = Rutgers. r is the correlation skill value and $MSSS$ is the percentage improvement in mean-square skill score over climatology. p is the probability that the observed skill value was obtained by random chance, with values shown in brackets. Bold type denotes skill values significant at less than 5% as determined by a Monte Carlo resampling test with 25,000 iterations.	91
3.4	Matrix of cross correlations between predictors for all three assessment periods. Values are absolute Pearson product-moment correlation coefficients. Bold type denotes values significant at less than 5% as determined by a Monte Carlo resampling test with 25,000 iterations. Data set abbreviations are as in Table 3.3.	96

3.5	As Table 3.3 except for lag-1 autocorrelation ($r[\text{lag-1}]$) of NAO_{DJF} indices and skill values from cross-validated hindcasts using the three NAO_{DJF} indices as predictors.	99
4.1	Linear trends in annual and seasonal SCA indices 1972-2002 estimated using an ordinary least-squares regression fit. Trends are the percentage change in seasonal SCA per decade. Bold type denotes significant ($p < 0.05$) trends as determined by a Mann-Kendall test.	121
4.2	Rank correlation coefficients between detrended annual and seasonal SCA indices for North America (NA), Eurasia (EU) and NH 1972-2002. Bold type denotes significant ($p < 0.05$) correlations (corrected for serial correlation).	124
4.3	Rank correlations between summer (JJA) to winter (DJF) SCA and DJF to JJA SCA indices for observations and ECHAM4 1972-2002 over NA, EU and NH.	128
4.4	Pearson pattern correlations between seasonal mean net surface radiative (NET) and sensible plus latent (TUR) fluxes for observations and ECHAM4 1972-2002 based on high minus low EU and NA SCA terciles. Left side shows values for EU, right side shows values for NA.	136
5.1	Skill values from ECHAM4 cross-validated NAO_{DJF} hindcasts for 1937-69 (top) and 1972-2002 (bottom) using JA mean NH SCA. r is the correlation skill value and $MSSS$ is the percentage improvement in mean-square skill score over climatology. p is the probability that the observed skill value was obtained by random chance, with values shown in brackets. Skill significance levels determined by a Monte Carlo resampling test with 25,000 iterations. All data are detrended.	151
5.2	Rank correlation coefficients between contemporaneous NH SCA and ΔT and ΔT_{EG} 1972-2002 for JJ, JA and JJA seasonal means. Bold type denotes significant ($p < 0.05$) correlations (corrected for serial correlation). RAW denotes data with trends included, while DET denotes data with a linear trend removed.	156

List of Figures

1.1	Schematic showing the regional climatic impacts of the positive (left) and negative (right) phases of NAO_{DJF} . Surfaces mark SSTs and sea ice extension, arrows show the flow systems in ocean, atmosphere and rivers, blue and red lines indicate near-surface sea-level pressures and white rectangles describe characteristic climate conditions or important processes. After Wanner et al. (2001).	21
1.2	Definitions of predictor period, forecast lead time and predictand period. After WMO (2002) and Lloyd-Hughes (2002).	23
1.3	The temperature, mean sea-level pressure (MSLP) and precipitation effects of the positive (panels a and b) and negative (panels c and d) phases of the winter NAO. Panels a and c: contours represent DJF MSLP anomalies (contour interval is 2.5 hPa and negative contours are dashed), shading represents DJF 2 m NCEP/NCAR air temperature anomalies (shaded at 0.25, 0.5, 1, 2 and 3°C where blues are negative and reds are positive). Panels b and d: shading represents anomalies in DJF mean CRU 0.5° rainfall accumulation (shading interval is 5mm). Positive and negative NAO phases relate to mean of upper and lower NAO_{DJF} terciles 1951-2001. Anomalies are calculated with respect to the 1950-2001 mean. MSLP data are area weighted by the cosine of latitude.	37
1.4	Correlations between monthly geopotential height anomalies and the NAO (left panels) and AO (right panels), respectively, for all months November to April 1958-97. Panels a and d show maps of the correlation at 300 hPa; panels b and e show meridional cross-section at 40°W; panels c and f show meridional cross-section at 10°E. Geopotential height data are from the NCEP/NCAR reanalysis, AO data are from Thompson and Wallace (1998) and NAO is based on index of Hurrell (1995). After Wanner et al. (2001).	39

-
- 1.5 Cross-validated NAO_{DJF} hindcasts 1972-2001 from dynamical and empirical SCF models. Line colours denote: observations (CRU NAO_{DJF}) (grey), DEMETER ensemble mean (red), May SST (SVD) model (blue) and JJ NH Snow cover model (yellow). The NAO_{DJF} values are standardised with respect to the 1972-2001 mean. 44
- 1.6 The correlation between lagged NH snow cover and three NAO_{DJF} indices 1972-2002 for bi-monthly snow cover periods ranging from JF (January-February) to ND (November-December). Negative correlation coefficients calculated with detrended time series are plotted. Dashed lines denote confidence levels corrected for serial correlation. After Saunders et al. (2003). 46
- 1.7 The energy budget above a snow covered surface. White circles denote transfers of energy: Q_e is latent heat flux through condensation or sublimation, Q_h is sensible heat flux through convection, Q_p is heat advected by rain falling on snow, “)Q” is energy exchange within snow pack and snow melt, Q_g is soil heat flux through conduction. “K” is short-wave solar radiation, “L” is long-wave infrared radiation, “\” denotes energy directed towards the surface and “[” denotes energy directed away from the surface. (Courtesy of D. Cline). 48
- 1.8 Time series of CRU NAO_{DJF} (black line), October mean Eurasian snow cover (blue) and JJ mean NH snow cover (red). Snow cover series are multiplied by -1 for display purposes. All time series are detrended and standardised with respect to the 1972-2001 mean. . . 51
- 2.1 Comparison of global reconstructed MSLP data sets. Lines show annual spatial Pearson correlations (60°S - 60°N) between the Smith and Reynolds (2004b) data (“Recon”), Basnett and Parker (1997) (thin solid line, “BP97”), Kaplan et al. (2000) (dashed line, “KEA”) and NCEP/NCAR reanalysis (thick solid line, “Rean”). After Smith and Reynolds (2004b). 61
-

2.2	DJF mean 2 m air temperature anomalies from CRUTEM2 (a and b) and NCEP/NCAR reanalysis (c and d) data. Data are for 1963 (top panels) and 1989 (bottom panels), which correspond, respectively, to extreme negative and positive NAO phases. Missing data in CRUTEM2 are white space. Colour levels are at 0.5, 1, 2, 3 and $> 3^{\circ}\text{C}$, where blues indicate negative and reds indicate positive temperature anomalies.	69
2.3	A comparison of global SST coverage for a week in December 1999 based on: <i>In situ</i> observations (top), satellite observations (middle) and the NCEP OI analysis model (bottom). The colour bar indicates the SST anomaly in $^{\circ}\text{C}$. (Courtesy of R. W. Reynolds).	71
2.4	Smoothed annual SST anomalies averaged over (a) Gulf Stream region (35°N - 45°N , 70°W - 50°W); (b) Kuroshio region (30°N - 40°N , 125°E - 160°E); (c) Greenland region (50°N - 70°N , 70°W - 30°W); (d) Baltic Sea; (e) Southern Ocean south of 50°S ; (f) Pacific Niño 3.4 region (5°S - 5°N , 170°W - 120°W). Anomalies are calculated with respect to the 1961-90 mean and only ice-free grid boxes with data are used. A 21-point binomial (near-decadal) filter was applied to annual anomalies. Thin black curves are unsmoothed HadISST1 annual anomalies. “HadMAT1” refers to the Hadley Centre night time marine air temperature data set, “OI.v2” refers to Reynolds et al. (2002), “GISST3.0” refers to an updated version of Rayner et al. (1996), “HadSST” refers to Jones et al. (2001) and “Kaplan et al.” refers to Kaplan et al. (1998). After Rayner et al. (2003). Figure is © Crown Copyright 2003.	76
3.1	Temporal evolution of the (a) CRU, (b) Hurrell and (c) MSLP PC1 NAO_{DJF} indices 1900-2001 in standardised units. Solid line indicates linear trend as determined by a least squares fit.	82

3.2	The correlation pattern significance between detrended time series of June-July Northern Hemisphere snow cover extent and gridded June-July 2 m air temperature. Significances are corrected for serial correlation with lags out to 15 years included. Colour scale denotes where correlation is positive (reds) or negative (blues). After Saunders et al. (2003).	85
3.3	Correlation coefficient between the lagged seasonal ΔT temperature index and the upcoming CRU NAO_{DJF} index. Top panel shows correlations for the CRUTEM2 ΔT index and bottom panel is for NCEP/NCAR data. Dark lines denote correlation coefficient and faint lines denote 5% significance level corrected for serial correlation.	86
3.4	Correlation coefficients between the JJ and MJJAS ΔT indices and the upcoming CRU NAO_{DJF} index for running 30-year windows commencing 1900 to 1971. Faint dashed line indicates 5% significance level corrected for serial correlation.	95
3.5	Composite vertical cross section of zonally averaged zonal wind anomalies based on high minus low terciles of June-July ΔT index 1950-2001. Zonal averages are calculated over (a) Eurasia, 25°E-70°E, (b) North America, 120°W-90°W and (c) North Atlantic, 50°W-20°W. Data are dimensionless standardised anomalies. Contour interval is 0.3 and dashed contours denote negative anomalies. Shaded areas denote values significant at less than 5% as determined by a Student's t -test.	101
3.6	Composite of (a) vertically averaged (925-200 hPa) North Atlantic sector zonal wind anomalies, (b) North Atlantic sector mean sea-level pressure anomalies (MSLP) and (c) North Atlantic SST anomalies based on high minus low terciles of June-July ΔT index. Contour interval is (a) 1 m s ⁻¹ , (b) 1 hPa and (c) 0.2°C. Dashed contours denote negative anomalies and zero contour is labelled. Shaded areas denote values significant at less than 5% as determined by a Student's t -test.	103

3.7	As Figure 3.6 except for winter (DJF) seasonal means (a) preceding and (b) following high minus low June-July ΔT terciles.	104
3.8	Area average of absolute standardised monthly mean anomalies of SST (solid line, dimensionless) and vertically averaged (925-200 hPa) zonal wind (dashed line, dimensionless) associated with high minus low terciles of June-July ΔT index. Values represent area average over the region [20°N-65°N, 100°W-0°E] for SST and the region [40°N-65°N, 120°W-40°E] for zonal wind.	106
4.1	Spatial distribution of seasonal mean snow cover extent for winter (DJF), spring (MAM), summer (JJA) and autumn (SON) for observations (left), ECHAM4 (centre) and observations minus ECHAM4 (right). Colours indicate mean fractional coverage. Shading interval is 0.1, with values <0.1 not coloured.	119
4.2	Percentage difference in monthly mean fractional SCA between observations and ECHAM4 1972-2002. Values are plotted for land area over Eurasia (EU), North America (NA) and the Northern Hemisphere north of 20°N (NH).	120
4.3	Time series of detrended seasonal mean SCA anomalies 1972-2002 for NA (left), EU (centre) and NH (right). Observed values are in blue and ECHAM4 values in red. Data are percentage SCA anomalies. Rank correlation coefficients are displayed.	123
4.4	Wavelet power spectra of seasonal mean NH SCA 1972-2002 for observations (left) and ECHAM4 (right). y -axis shows period in years and cone of influence is indicated by cross-hatching. Colour shading indicates wavelet power at levels $S/3$, $S/2$, S , $2S$ and $3S$, where S is the standard deviation of the seasonal wavelet power. Black contour denotes significant ($p < 0.05$) power.	125
4.5	Autocorrelation function out to a time lag of 10 years for seasonal NA, EU and NH SCA indices 1972-2002. Blue line is observed SCA and red line is ECHAM4 SCA.	127

4.6	Zonally averaged (50°E-100°E) Eurasian atmospheric temperature on pressure levels for high minus low EU SCA terciles for observations (NCEP/NCAR reanalysis, top panels) and ECHAM4 (bottom panels) at latitudes 20°N-80°N. Data are temperature anomalies and contours are plotted at 0.25, 0.5, 1, 2 and 4°C. Dashed contours indicate negative anomalies and the thick solid contour denotes the zero line. Shading denotes significant ($p < 0.05$) differences as determined by a Student's t -test.	130
4.7	As Figure 4.6 except for zonally averaged (120°W-90°W) NA atmospheric temperature, NA SCA terciles and contours are plotted at 0.5, 1, 2, 3 and 4°C.	132
4.8	Observed (NCEP/NCAR reanalysis, left) and ECHAM4 (right) seasonal mean surface net radiative flux anomalies (Wm^{-2}) over EU for high minus low EU SCA terciles 1972-2002. Colours indicate positive (reds, surface losing energy) or negative (blues, surface gaining energy) fluxes.	134
4.9	As Figure 4.8 except for sensible plus latent heat fluxes.	135
4.10	As Figure 4.8 except for NA net surface radiative fluxes and SCA anomalies.	137
4.11	As Figure 4.8 except for NA sensible plus latent heat fluxes and SCA anomalies.	138
5.1	Temporal evolution of the ECHAM4 NAO indices. The indices are from ECHAM4 MSLP data and calculated using the same methods and data regions employed in constructing the observed indices shown in Figure 3.1, which are (a) CRU, (b) Hurrell and (c) MSLP PC1 NAO_{DJF} 1904-2002 in standardised units.	145

5.2	Wavelet power spectra of NAO_{DJF} indices 1904-2002. Plots show the observed (a) CRU, (b) ECHAM4 CRU, (c) ECHAM4 Hurrell and (d) ECHAM4 MSLP PC1 NAO_{DJF} indices. Cone of influence is indicated by cross-hatching. Colour shading indicates wavelet power at levels $S/3$, $S/2$, S , $2S$ and $3S$, where S is the standard deviation of the seasonal wavelet power. Black contour denotes significant ($p < 0.05$) power.	147
5.3	Lead/lag rank correlation matrices between monthly mean NH SCA and the NAO for observations (top), ECHAM4 CRU (centre) and ECHAM4 Hurrell (bottom). Contemporaneous correlations denoted by diagonal line. Elements left of diagonal denote SCA leading and right of diagonal denote NAO leading. Colour shading denotes statistical significance assuming 30 degrees of freedom at 0.01, 0.05 and 0.1 levels; blues indicate negative, while reds indicate positive correlation. Coefficients where $p > 0.1$ are not plotted. All data are detrended.	149
5.4	Lagged rank correlation between ECHAM4 bi-monthly mean NH SCA and the three NAO_{DJF} indices for (a) 1904-36, (b) 1937-69 and (c) 1972-2002. Dotted lines denote statistical significance ($p < 0.05$) corrected for serial correlation. All data are detrended.	152
5.5	Rank correlation coefficients between JJA mean NH 2 m air temperature and NH SCA for observations (NCEP/NCAR reanalysis, left) and ECHAM4 (right) 1972-2002.	154
5.6	JJA mean 500 hPa geopotential height anomalies for observations (top), ECHAM4 (centre) and observations minus ECHAM4 (bottom) associated with high minus low JJA NH SCA terciles 1972-2002. Contour interval is 5 dm for top and centre panels and 10 dm for bottom panel. Red contours denote positive values and blue contours denote negative values. Grey shaded areas denote differences significant at 5% as determined by a Student's t -test.	157

-
- 5.7 As Figure 5.6 except for JJA 850 hPa winds in observations (top), ECHAM4 (centre) and observations minus ECHAM4 (bottom). Vectors show monthly mean wind anomalies in ms^{-1} 159
- 5.8 Pearson pattern correlations between contemporaneous observed and ECHAM4 2 m air temperature composite means for high minus low terciles of Greenland temperature (solid line) and NH SCA (dashed line) 1972-2002. Greenland temperature is area average for region $[67^{\circ}\text{N}-72^{\circ}\text{N}, 52^{\circ}\text{W}-47^{\circ}\text{W}]$ 160
- 5.9 As Figure 3.6 except for ECHAM4 JJA mean (a) 850 hPa zonal wind, (b) MSLP and (c) SST anomalies associated with high minus low JJA NH SCA terciles 1972-2002. NH SCA index is multiplied by -1 to facilitate comparison with Figure 3.6. 162
- 5.10 As Figure 5.5 except for ECHAM4 1904-36 (left) and 1937-69 (right). 166
- 5.11 Pearson pattern correlation coefficients between fields of 2 m air temperature for pairs of adjacent months denoted by their first letters. Temperature fields are the composite mean associated with high minus low JJA NH SCA terciles. Line colours denote: observations 1972-2002 (black), ECHAM4 1972-2002 (blue), ECHAM4 1937-69 (orange) and ECHAM4 1904-36 (red). Dashed horizontal lines show pattern correlations at $r = 0.4$ and $r = 0.6$, respectively. 167

Acknowledgements

This work represents a large proportion of my adult life and I am indebted to many people for their help and influence along the way: The UK Natural Environmental Research Council for generously funding my research. St Paul Re for providing CASE sponsorship. Mum and dad for always being there. My brother for taking a call-centre job and finally letting me close the gap. Dr Mark Saunders for supervision. Ben Lloyd-Hughes for reviewing the snow work. Adam Lea and Fiona Parton for feedback. Alan Muir for turning remote computing into a reality. Abi Rymer for putting me up when home was far away. Matt Whillock for letting me win at table tennis (but never snooker). John Allen at the Jitsu Foundation. The data providers whose products are used in this thesis; particularly, the British Atmospheric Data Centre and the Deutsches Klimarechenzentrum. The good people of the Mullard Space Science Laboratory for providing a unique and special working environment. The housemates, friends and lovers who have kept it fun throughout. The musicians who recorded the soundtrack to it all. And finally, to the grandparents, who look on as we make all the same mistakes again.

Chapter 1

Introductory Material

1.1 Introduction

Winters in Europe are known to exhibit different climatic characteristics from one year to the next. However, it is also known that certain weather types persist for several winters in succession (Johansson et al. 1998, Wanner et al. 2001). Increasing research is focusing on understanding the causes of variations in UK and European winter weather on seasonal, interannual and decadal time scales. In particular, recent attention has focused on the seasonal forecasting of large-scale climate modes, which influence weather patterns on hemispheric spatial scales.

The North Atlantic Oscillation (NAO) is the dominant mode of winter climate variability over the North Atlantic (Walker and Bliss 1932, Barnston and Livezey 1987). The NAO is characterised by a dipole oscillation in North Atlantic sea-level pressure between the subpolar (Icelandic) low pressure and subtropical (Azores) high pressure regions (Saunders and Qian 2002). The NAO index represents the strength of the meridional sea-level pressure gradient between these regions. High values of the winter NAO index are linked to increased storminess over northern Europe, while low values are linked to increased storminess over southern Europe and the Mediterranean (Figure 1.1). Teleconnected atmospheric and oceanic effects associated with the winter NAO are also observed over a wide area surrounding the North Atlantic. Accurate and timely forecasts of the winter NAO index are therefore an important challenge for seasonal forecasters and current research is directed at improving the quality and lead time of NAO forecasts.

Seasonal climate forecasting aims to predict the mean state of the atmosphere

months or seasons in advance. Research over the last couple of decades has shown that, in many regions of the world, the seasonal climate is potentially predictable (Goddard et al. 2001). The aim of this thesis is to investigate the seasonal predictability and modelling of the winter NAO. The main objective is to use empirical models to determine the current level of available seasonal predictability. This work builds on previous forecasting attempts by presenting improved seasonal predictive models for the winter NAO.

This Chapter begins with a description of the background and process of seasonal climate forecasting (Section 1.2). A review and comparison of the two different methodologies employed in solution to the problem is presented. Section 1.3 outlines the need for accurate seasonal forecasts of the winter NAO and identifies potential users of the forecasts. Section 1.4 defines the first research topic in this thesis, which compares four empirical seasonal prediction models for the winter NAO. Section 1.5 outlines the second research topic, which examines the capability of a coupled general circulation model (GCM) to reproduce observed winter NAO predictability from lagged snow cover. The motivation for these research areas is outlined, along with the specific research questions that will be addressed by this work. A brief summary is presented in Section 1.6.

1.2 Seasonal climate forecasting

1.2.1 Physical basis

Seasonal climate forecasting (SCF) is a revolutionary technology aiming to predict the likely state of the climate several months ahead (Mathieu et al. 2004). SCF has been performed for over a century based on the identification of environmental indicators that precede shifts in climate (Goddard et al. 2001). However, SCF has received renewed interest during the past 20 years through the realisation that the El Niño Southern Oscillation (ENSO) phenomenon offers potential global seasonal predictability. The scientific basis for SCF lies in the lower boundary of the atmosphere (i.e., ocean and land), which acts as a climate pacemaker by modulating

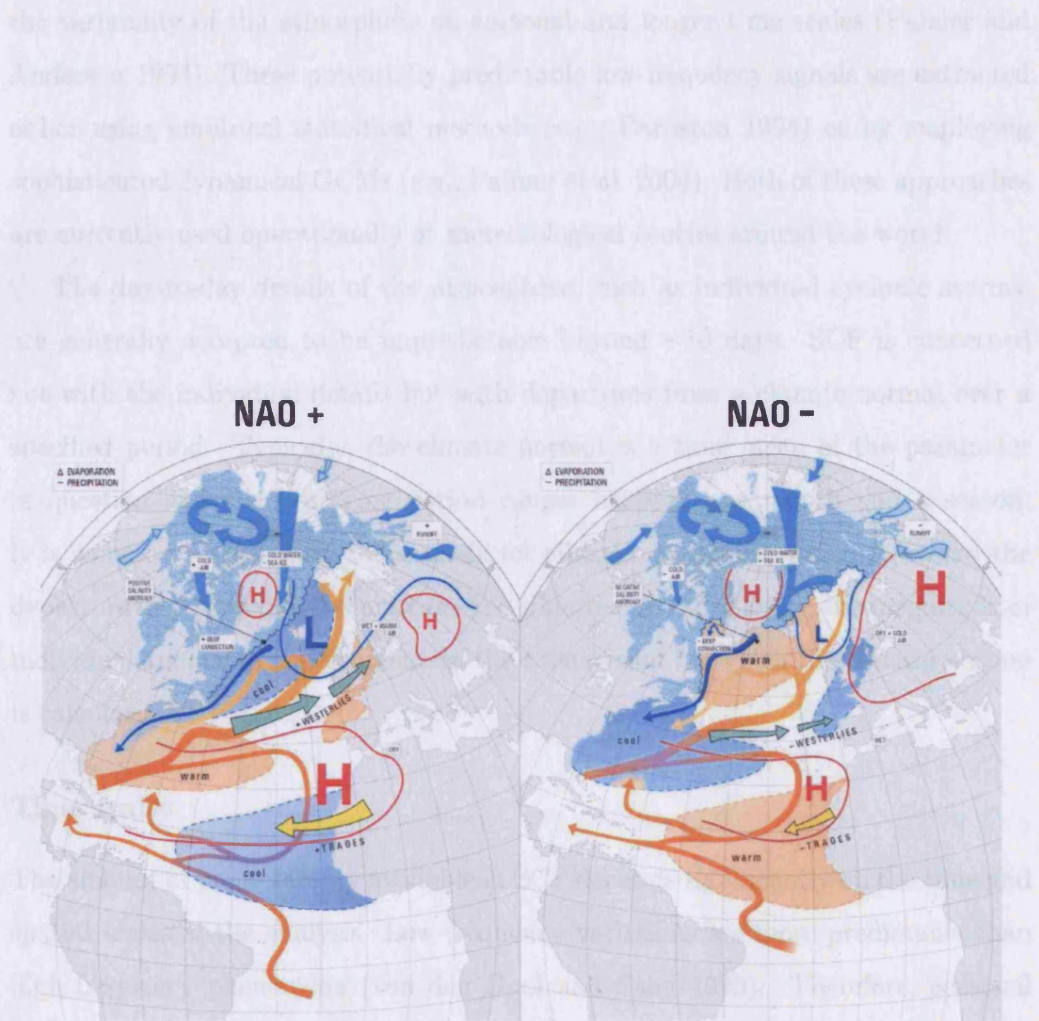


Figure 1.1: Schematic showing the regional climatic impacts of the positive (left) and negative (right) phases of NAO_{DJF} . Surfaces mark SSTs and sea ice extension, arrows show the flow systems in ocean, atmosphere and rivers, blue and red lines indicate near-surface sea-level pressures and white rectangles describe characteristic climate conditions or important processes. After Wanner et al. (2001).

the variability of the atmosphere on seasonal and longer time scales (Palmer and Anderson 1994). These potentially predictable low-frequency signals are extracted either using empirical statistical methods (e.g., Barnston 1994) or by employing sophisticated dynamical GCMs (e.g., Palmer et al. 2004). Both of these approaches are currently used operationally at meteorological centres around the world.

The day-to-day details of the atmosphere, such as individual cyclonic storms, are generally accepted to be unpredictable beyond ~ 10 days. SCF is concerned not with the individual details but with departures from a climate normal over a specified period. Typically, the climate normal is a time mean of the parameter in question and the averaging period ranges between one month and a season. It is common for forecasts to be made for climate anomalies, which represent the departure of an individual event from the climate normal. Clearly, the magnitude of individual anomalies is dependent on the time period for which the climate normal is calculated.

Time scales

The amount of predictability available in SCF depends significantly on the time and spatial scales of the analysis. Low frequency variations are more predictable than high frequency phenomena (van den Dool and Saha 1990). Therefore, seasonal forecasts are issued for a time mean period of predetermined length rather than for individual weather events. The definitions of the forecast lead time and the predictor and forecast periods are shown in Figure 1.2. Typical time scales for the predictor and validation periods involved in SCF range between a single month and a year. However, a season is considered the shortest averaging period that allows an acceptable signal-to-noise ratio and the longest averaging period that does not mix different times of the year too much (Johansson et al. 1998).

Forecast lead times vary from 0-3 months (e.g., Saunders and Qian 2002, Palmer et al. 2004) to as long as one year (e.g., van den Dool and Nap 1985). Clearly, a longer lead time is more valuable to an end user. However, the maximum useful lead time is dependent on the skill of the forecast at that lead. For example, given a

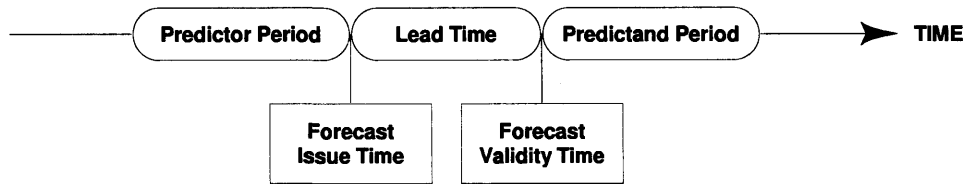


Figure 1.2: Definitions of predictor period, forecast lead time and predictand period. After WMO (2002) and Lloyd-Hughes (2002).

choice between a forecast with low skill 6 months ahead and a forecast with highly significant skill 3 months ahead, the end user would have to decide whether the extra 3 months lead time offsets the reduced accuracy of the prediction. In practice, forecasts are typically issued at a range of lead times and the skill increases closer to the validation time (e.g., Lloyd-Hughes et al. 2004).

Spatial scales

The spatial scale of the forecast region is also important in SCF. Lorenz (1969) presented theoretical arguments showing that climatic phenomena occurring on large spatial scales have a longer range of predictability than those on smaller spatial scales. This consideration means that seasonal forecasts tend to be made for climatic regions (e.g., eastern United States, northwest Europe etc). However, seasonal forecasts can also be made for point locations where a single weather station is situated in a particularly localised climate regime (e.g., at a coastal boundary) (van den Dool and Nap 1981).

On the global scale, most SCF predictability derives from the ENSO phenomenon. ENSO is the dominant mode of global climate variability on seasonal-to-interannual time scales and arises from coupled ocean-atmosphere interactions in the tropical Pacific Ocean (Bjerknes 1966, Wallace et al. 1998). Although centred on the Pacific, ENSO has major teleconnected impacts around the globe, which are thought to arise through the propagation of tropical Pacific wavelike disturbances (Trenberth et al. 1998). Mathieu et al. (2004) showed that SCF skill is observed in the North Atlantic and European (NAE) sector during ENSO events. However, the NAE sector relationship with ENSO is complex and these authors

stated that SCF systems must focus on local conditions as well as those over the tropical Pacific Ocean. This point is further emphasised by the obvious recognition that Atlantic conditions are likely to be still more important during non-ENSO years. Graham et al. (2000) found that ENSO-related improvements in predictability over the Northern Hemisphere apply predominantly to North America and that in non-ENSO years the levels of skill are similar between locations.

Aside from ENSO, hemispheric and regional scale SCF predictability is generally concerned with the influence of patterns of lagged global and regional sea-surface temperature (SST) (Lorenz 1969). Tropical Atlantic SSTs have been shown to influence seasonal Atlantic hurricane activity (Saunders and Harris 1997, Goldenberg et al. 2001) and interannual rainfall variability from Brazil to West Africa (e.g., Hastenrath 1984, Enfield 1996). Extratropical North Atlantic SSTs have also been shown to play a significant role in SCF on hemispheric scales (Radcliffe and Murray 1970, Colman and Davey 1999, Rodwell and Folland 2002, Saunders and Qian 2002).

At smaller spatial scales in the extratropics, two SCF mechanisms have been identified based on direct and indirect climatic effects. The first mechanism is the direct persistence in atmospheric conditions caused by proximity to water bodies, which occurs on small spatial scales of the order ~ 200 km (van den Dool and Nap 1981). The second mechanism concerns the indirect large-scale persistence of the background atmospheric circulation. Extratropical synoptic weather systems have a dominant influence on climate at this larger spatial scale and therefore predictability is assumed to be lower than with the first mechanism (van den Dool 1983). However, in extratropical regions, for accurate SCF one has to combine the two effects (van den Dool and Nap 1981).

Verification

Seasonal forecasts only have value if they are accurate above and beyond what could be expected by random chance. The accuracy of SCF systems is commonly referred to as the forecast 'skill' and several different 'skill scores' have been developed to

quantify their value. The main purpose of skill scores is to chart the variation in the success of a given series of forecasts produced in a standard format and checked in a standard objective way (Gilchrist 1986). Skill scores also enable the accuracy of forecasts produced by different models and using different climatic variables to be compared. Skill is considerably more difficult to assess in SCF than in short- or medium-range weather forecasting because of the increased time scale of the forecast, and hence the reduced frequency with which forecasts can be made and evaluated (Zwiers and von Storch 2004).

Seasonal forecasts can be either probabilistic or deterministic and different skill scores are used for each type. Probabilistic forecasts provide probabilities of the occurrence or non-occurrence of an event or a set of fully inclusive events (WMO 2002). Some examples of verification measures are the Brier Score (Brier 1950), Relative Operating Characteristic (Palmer et al. 2000) and the Rank Probability Score (Wilks 1995). Deterministic forecasts provide a single expected value for the forecast variable (e.g., a predicted value of monthly mean temperature) (WMO 2002). They are typically assessed against a prediction of the climatological mean and verified using the Pearson or rank correlation, or mean-squared skill scores (Wilks 1995, WMO 2002).

1.2.2 Dynamical SCF

Having discussed the physical basis and the different time and spatial scales involved in SCF, we now examine the contrasting methodologies employed in solution to the problem. These methodologies are to use either dynamical or empirical climate models, respectively.

Dynamical SCF for a month or season ahead is distinguished from short- and medium-range dynamical forecasting by being concerned with a period that extends beyond the limit of predictability for individual weather systems. ‘Limit’ in this sense refers to the time taken in a numerical model simulation for small errors consistent with practical observational inaccuracies in the initial atmospheric state, to grow to saturation values (Gilchrist 1986). Palmer and Anderson (1994) stated

that for SCF it is not sufficient to simply expand the length of integration of a numerical weather prediction (NWP) model; the model must be coupled to the lower boundaries, most importantly to the oceans. Uncertainties in the initial conditions of these integrations cause random errors to erode forecast skill. Sensitivity to initial conditions may be filtered, to a degree, through use of the ensemble mean of a series of model integrations, each run with slightly different initial conditions (e.g., SST fields for a series of consecutive days). Ensemble meaning has proved relatively successful in reducing initial condition uncertainty in SCF (e.g., Lin et al. 2004). Such GCM ensemble integrations, forced with sufficiently accurate observed SSTs, represent an upper limit on the potential predictability available from coupled ocean-atmosphere GCMs (Branković and Palmer 2000).

While SCF ensembles from a single GCM reduce initial condition uncertainty, recent projects have sought to reduce the effects of model formulation error by employing an ensemble of predictions from different GCMs. The scientific basis for these multi-model ensembles is that each GCM formulation is different and the combined ensemble therefore explores forecast uncertainty due to model formulation (Krishnamurti et al. 1999). One of the earliest examples of this approach was the PRediction Of climate Variations On Seasonal to interannual Time scales (PROVOST) project (Palmer et al. 2000, Branković and Palmer 2000, Graham et al. 2000). PROVOST showed that, using four atmospheric GCMs, the multi-model skill exceeds that of the individual models. Effectively, the multi-model acts as a filter for the strongest model in the ensemble (Graham et al. 2000). More recently, the DEMETER project (Palmer et al. 2004) used seven independent global coupled GCMs to produce multi-model ensemble hindcasts 1980-2001. DEMETER hindcast skill appears in general to be an improvement on that seen in PROVOST. However, comparisons are difficult because different assessment time periods and verification times are employed.

A major limitation of the multi-model ensemble approach is the influence of systematic errors in the GCM climate evolutions. Each GCM in the ensemble exhibits a particular pattern of errors in its output. The errors in a single model can be of a similar magnitude to the signal being predicted (Branković and Palmer 2000).

The multi-model climatology is used to validate the hindcast skill, which means the skill assessment may be biased. A procedure for correcting bias in multi-model ensembles has been outlined (Palmer et al. 2000, 2004). This involves applying a linear correction to the output from each GCM, by expressing the prediction as an anomaly relative to the GCM climatology. The resulting skill values are therefore a good approximation to ‘true’ independent hindcast skill. However, model biases are not necessarily linear, and this correction could be improved (F. Doblas-Reyes personal communication, January 2005).

The multi-model approach has shown success in reducing uncertainties from initial conditions and model formulation. However, it cannot be considered the final solution to dynamical SCF (Palmer et al. 2004). This is because even the most sophisticated coupled GCMs have inherent deficiencies and these are often shared between models (e.g., representation of subgrid-scale processes such as convection). Palmer et al. (2004) stated that the DEMETER results have motivated a more theoretical approach to representing model uncertainty using stochastic-dynamic subgrid models. The future of dynamical SCF will also rely heavily on the expected improvements in the physics of coupled GCMs. In particular, improved representation of land surface coupling may lead to increases in SCF predictability (e.g., Frei et al. 2003, Gong et al. 2004).

1.2.3 Empirical SCF

Lorenz (1970) stated that ‘weather is what you get; while climate is what you expect’. Empirical SCF follows this principle, with forecast models formulated based on the way the climate system has evolved in the past. Statistical methods are employed on historical observational data sets to identify relationships between a predictand variable and one or more predictor variables. The major benefit of empirical, compared to dynamical, SCF is the relative reduction in computing power required to produce a forecast. This makes empirical SCF feasible at meteorological centres without access to supercomputers (e.g., in the developing world). Furthermore, the expanding global repository of historical climate observations means that

empirical modellers are able to examine a huge number of combinations of predictive relationships.

The most reliable SCF tools are presently statistical (Zwiers and von Storch 2004). Powerful new statistical techniques such as neural networks (e.g., Hsieh and Tang 1998) and cluster analysis (e.g., Mimmack et al. 2001) have been employed in an effort to improve empirical models. However, Zwiers and von Storch (2004) state that much of the work that has had a large impact on climate research has used relatively simple techniques. In this thesis, we focus on simple statistical techniques such as linear regression (Wilks 1995) and also more sophisticated data reduction techniques such as principal components analysis (PCA) (von Storch and Zwiers 1999).

Empirical methods also have a number of disadvantages and limitations. First, great care is required to find relationships that are based on the physics of the climate system. A statistically significant correlation between two variables does not imply that those variables are physically linked. Many spurious associations have been identified in this way both in climatology and other disciplines (Zwiers and von Storch 2004). In any global map of correlation with a sample size of $N = 30$, on average 5% of the area will show coefficients greater than $|r| = 0.36$ in the absence of any 'real' significant correlation. Furthermore, gridded climate fields contain relatively few spatial degrees of freedom because individual grid points are not independent. This can lead to erroneous interpretation of empirical results (Livezey and Chen 1983). Careful and detailed investigation using all available data must be undertaken to ascertain cause and effect. Often, the only credible way in which to discriminate between statistically 'significant' features and truly dynamical features of the system is dynamical plausibility and reproducibility in detailed dynamical models (Zwiers and von Storch 2004).

Second, the stability of significant predictive links with time is an important test for empirical models. Predictive relationships are seldom stationary when assessed over periods of decades and longer. One example is the predictability of the Indian monsoon, where forecast skill has been shown to vary on inter-decadal time scales associated with the ENSO phenomenon (Sahai et al. 2000). Nonstationarity in pre-

Empirical		Dynamical	
Advantages	Disadvantages	Advantages	Disadvantages
Uses all historical data	Non-linear effects usually not modelled	Global coverage	Spatial resolution
Identifies precursive relationships	Physical mechanisms may not be elucidated	Physical mechanisms can be established	Not as well developed as statistics
Useful predictions out to ~ 9 months	Difficulty predicting for new regimes	Regime changes can be modelled	Unclear that skill will exceed that of statistics
Reduced CPU/cost required	Restricted number of predictors	More direct treatment of uncertainties through ensembles	Increased CPU/cost required

Table 1.1: Summary comparison of empirical and dynamical methods used for SCF. After Lloyd-Hughes (2002).

dictive relationships can also occur due to creep or sudden changes in instrumental measurement systems for particular climate variables (Zwiers and von Storch 2004). These changes are discussed in greater detail in Chapter 2.

Third, empirical models are reliant on the temporal resolution of the input data. This means that if monthly data are employed, only climatic processes acting consistently on time scales ≥ 1 month will be captured. Given the dynamic nature of the extratropical atmosphere, on these time scales many important processes will either not be observed or ‘smoothed out’ by averaging. These processes can be either subgrid-scale (e.g., sea spray or small convective cells) or temporally limited (e.g., diurnal radiative effects or synoptic weather systems). The problems discussed above show that consistent operational verification over years and, preferably, decades is required to truly evaluate the utility of empirical SCF models.

1.2.4 Dynamical versus empirical SCF

The main advantages and disadvantages of empirical and dynamical SCF are summarised in Table 1.1. Gilchrist (1986) stated that empirical models derived using

statistical methods may provide a better way of investigating extended-range predictability than using GCMs. While statistical methods are currently prominent, the widely accepted goal is to produce reliable dynamical SCF systems (Zwiers and von Storch 2004). The argument often put forward in favour of dynamical SCF is that GCMs have a clear development path, which will lead to improved forecasts in time (e.g., Palmer and Anderson 1994, Goddard et al. 2001). However, past experience has shown that improving GCMs has not been straightforward and, consequently, progress in dynamical SCF systems has been relatively slow (van Oldenborgh et al. 2003).

Traditionally, climate modellers have fallen into either the empirical or the dynamical camp but seldom both. However, the opposition of dynamical and statistical methods is a false antithesis and the possibility exists of combining the benefits of both methods (van Oldenborgh et al. 2003). Indeed, most dynamical SCF systems are presently ‘two tier’ (e.g., Kanamitsu et al. 2002, Lin et al. 2004). This means that the lower boundary conditions (usually SSTs) are forecast using empirical methods and are then employed to force a dynamical model. Another recent example is described by Barnston et al. (2003), who combined a fully coupled GCM with a statistical model to produce a more skilful operational SCF model.

There are several important factors to consider when selecting an appropriate SCF methodology. First, dynamical SCF can only be conducted effectively with supercomputing resources (including technical support). Second, empirical and dynamical SCF make different demands on data sources. Empirical SCF requires long, homogeneous climate records for specific locations, while dynamical SCF often requires daily, global fields to initialise GCM simulations. Third, the time required to conduct dynamical SCF experiments is much greater than for empirical modelling. Fourth, the methodology should be selected appropriate to the level of expertise at a particular research centre. These factors, coupled with the previous discussion, suggests that empirical methods are the preferred method for the purposes of this study.

1.3 Potential benefits of SCF

As part of the motivation for undertaking this study, it is important to assess the potential benefits and identify end users of the timely and accurate winter NAO forecasts. A seasonal forecast is only of more than academic use if there is an end user who attaches value to that forecast. Furthermore, forecast quality is ultimately in the perception of the beholder and not just in the evaluation of the forecaster (Hartmann et al. 2002). A numerically accurate forecast is therefore of little use unless the information is easily interpreted by end users.

In Section 1.3.1 we define the need for, and potential benefits from, accurate and timely forecasts of the winter NAO. Second, we identify some of the end users who could benefit from these forecasts. The users fall into two broad categories, namely, industrial and societal. These categories are based on the different methods of payment for the forecast output. Industrial users are expected to fund privately the forecasts that they use because the data will be employed for financial gain. By contrast, with societal users, the cost is normally paid by a government or national meteorological service because the information is of benefit to society. These user groups are examined in more detail in Section 1.3.2.

1.3.1 Benefits of seasonal forecasting for the winter NAO

The winter NAO is a large-scale climatic mode that influences seasonal weather patterns from November to April over a large region surrounding the North Atlantic (Hurrell 1995). The climatic impacts of changes in the phase of the winter NAO are observed in temperature, precipitation and storminess. Consequently, agricultural harvests, water management, energy supply and demand, and yields from fisheries, among many other things, are directly affected by the NAO (Hurrell et al. 2003). This means that there exists a large potential demand for accurate and timely forecasts of the winter NAO.

Many studies have investigated the economic, ecological and societal impacts of NAO variability. Here, we present the findings of just a few to demonstrate the benefits that could be gained from skilful winter NAO forecasts. The largest

economic stake in NAO-related risk appears to exist in the energy sector. Hurrell et al. (2003) described how Norway and Sweden began to trade energy in the early 1990s based on winter NAO forecasts. The two countries are ideally suited to this arrangement because their climates are affected differently by the NAO and their energy production is distributed differently between hydro, nuclear and fossil fuel power. Consequently, during negative NAO phases when precipitation is low, Norway tends to supplement its hydropower production by purchasing nuclear power from Sweden. During positive NAO winters of high precipitation over Norway, Sweden finds it cost effective to purchase hydropower from Norway.

Trigo et al. (2004) found that interannual variability in winter precipitation and river flow on the Iberian peninsula are related significantly to changes in the NAO index one month in advance. It was shown that accurate forecasting of the NAO could yield major benefits to the local hydroelectric power industry, which accounts for ~20% of the total annual electricity production. However, these authors highlighted that winter NAO predictability is nonstationary and varies on decadal time scales. Winter NAO predictability is known to be higher since 1950 than before 1950 (e.g., Rodwell and Folland 2002), although such variability may be a natural feature of the climate system (e.g., Osborn et al. 1999).

NAO forecasts can also have political and humanitarian implications for a country or region. Cullen and deMencol (2000) studied the connection between the NAO and stream flow of the Euphrates and Tigris rivers. Water supply shortages and surpluses are a major issue in this region for irrigation farming. Decreases in rainfall over the Middle East, associated with the positive trend in the NAO index since the 1970s, have had catastrophic effects on crop yields. Consequently, high level political disputes have arisen because of water withdrawals from the rivers between Turkey, where most of the rain falls, and Syria, situated downstream. Accurate NAO predictions would therefore be an important tool for water resource management in this region and may help to diffuse potential disputes before they arise.

1.3.2 User groups

Industrial

Traditionally, weather-related risk was considered only to be significant for companies whose output was specifically weather-dependent (e.g., energy providers, ice-cream manufacturers etc.). However, the weather risk market has expanded such that the energy, banking, commodity trading and insurance/reinsurance sectors are all now represented (Clemmons 2002). Individual companies can act to offset their weather-related exposure by using SCF. The increasing awareness of the potential industrial benefits to be gained from SCF has led to the creation of weather-specific financial markets in London and Chicago. These markets are driven by companies and investors hedging their assets against weather predictions for a range of time scales from days to seasons. In this case, the forecasts themselves are potentially worth a huge amount of money and the more accurate and timely a forecast system is, the greater its financial value.

A seasonal forecast delivered to a company or organisation is likely to be tailored to their specific needs in terms of location, time period and presentation. However, the accuracy of the forecast can only be as good as the model from which the data are derived. In the UK, many private weather consultancy firms currently issue forecasts to companies who pay significant amounts for the service. However, it is common for weather consultancy firms simply to amalgamate freely available data from several national meteorological services in producing the forecast product. This process can be thought of as a simple ‘ensemble’ prediction, adding interpretation and presentation but no additional climatological information to the forecast product.

Another major limitation in industrial SCF is the cost of the data used in producing the forecast and the forecast output data itself. Many SCF research projects have made their output data freely available to researchers for analysis (e.g., DEMETER). Problems may arise when these data are used not for research purposes but for financial gain. The data then derive significant value for the intellectual and resource investment in their creation. The costs involved in data

preparation and processing have the potential to render industrial SCF financially unviable if forecast skill is not high enough to warrant such high levels of investment.

Societal

A forecast made for society is likely to show lower levels of detail than a forecast made for a specific organisation or company. There are several intuitive reasons for this assertion. First, the number of end users is far greater and therefore it is simply impossible for a single meteorological service or seasonal forecast group to produce a point and time specific forecast for every person in the country. Typically, forecasts are published for cities or regions and cover periods of weeks, months and seasons (e.g., UKMO 2004). Second, the number of different purposes to which the end users will put the forecast information is almost as large as the number of users. Therefore, the information content must be general enough to have some value to the majority of users. Third, national meteorological services are becoming increasingly reliant on private, rather than government, investment to fund their operations. In terms of SCF, this means that a scaled-down forecast will be made freely available to the public, while private companies can pay for a more detailed and informative forecast if they so wish.

Societal users include health service providers who may need to plan 1-3 months ahead to purchase a certain quantity of influenza vaccinations. Prolonged anomalous cold weather during winter is associated with 'excess winter mortality' caused by significant increases in respiratory and cardiovascular diseases (Mercer 2003). Conversely, seasonal prediction of anomalous warm periods is also of critical importance, as seen in France during summer 2003, where more than 10,000 deaths resulted from a summer heat wave (Levinson and Waple 2004). Individual societal end users potentially include families wishing to plan holidays and gardeners interested in when and what to plant on their allotments next season.

There is a common perception both in the meteorological community and in society that seasonal forecasts are unreliable. This uncertainty about the accuracy of seasonal forecasts precludes users from making more effective use of them (Hart-

mann et al. 2002). The problem seems to stem from poor communication of the available skill and, in particular, the reliability of that skill. User groups must be informed of problems such as seasonality and nonstationarity in SCF models. Decision makers need to know when forecasts are not reliable enough for their purposes (Hartmann et al. 2002).

1.4 Seasonal predictability of the winter North Atlantic Oscillation

1.4.1 Definition and background

The overall theme of this thesis is to investigate the seasonal predictability of the winter North Atlantic Oscillation (NAO). This Section outlines the first research topic for this study, which compares the hindcast skill achieved using four published empirical winter NAO predictors. The NAO is defined as a meridional redistribution of atmospheric mass between the subpolar and the subtropical North Atlantic (Hurrell et al. 2003) and is the dominant mode of boreal winter climate variability over the North Atlantic sector (Walker and Bliss 1932, Barnston and Livezey 1987). The NAO is linked strongly to patterns of winter temperature, precipitation and storminess over the whole North Atlantic sector (Hurrell 1995, Trigo et al. 2002). Johansson et al. (1998) stated that the amplitude and time scale of the NAO are also large enough to drive persistent conditions over much of northern Europe and the North Atlantic sector for several winters in succession. Accurate and timely forecasts of the winter NAO are therefore an important challenge for seasonal forecasters and increasing research is focusing on improving forecast quality and lead time.

The term North Atlantic Oscillation was first used by Walker (1924) to describe a mode of climate variability that dominated world weather. Walker and Bliss (1932) were among the first to construct an NAO index, which related to temperature and pressure anomalies over northern Europe. More recently, Hurrell (1995) and Jones et al. (1997) constructed indices to represent the NAO based on localised pressure

differences between stations in Iceland and the Azores or Gibraltar, respectively. High NAO phases represent lower pressures over Iceland, while low NAO phases indicate a reduced meridional pressure gradient over the North Atlantic. The NAO can also be represented by the leading component of variance in North Atlantic sector gridded sea-level pressure fields (e.g., Hurrell 1995). The individual NAO indices used in this study, and methods employed in their construction, are discussed in detail in Section 2.3.

Figure 1.3 shows winter North Atlantic mean sea-level pressure (MSLP), 2 m air temperature and precipitation anomalies associated with positive and negative winter NAO phases 1950-2001. The positive phase (panels a and b) shows lower MSLP over Iceland and increased MSLP over western Europe and the North Atlantic associated with a northward shift of the North Atlantic storm track. The temperature pattern is consistent with increased warm advection over northern Europe from cyclonic storms. The negative phase (panels c and d) shows the inverse pattern in MSLP and temperature. The precipitation pattern in the positive phase results in increased accumulation over the UK, northwest Europe, Scandinavia and western Russia. During the negative phase, precipitation is increased across much of southern Europe and the Mediterranean region.

1.4.2 The winter season

Some justification of the narrowing of the research to the winter season is required. Furthermore, the term ‘winter season’ itself requires precise definition. The decision to focus this research on the winter season is taken for several reasons. First, the Northern Hemisphere (NH) winter atmosphere is dominated by large-scale modes of climate variability (e.g, the NAO) (Barnston and Livezey 1987). These modes are easily detected in observations and GCM data above the background climatic noise. Climate modes are assumed to be more predictable than individual weather events because they vary on time scales of weeks and months rather than hours and days. The reasons for this are discussed further in Section 1.4.4. Second, the influence of ENSO is known to be strongest in winter, which may offer increased

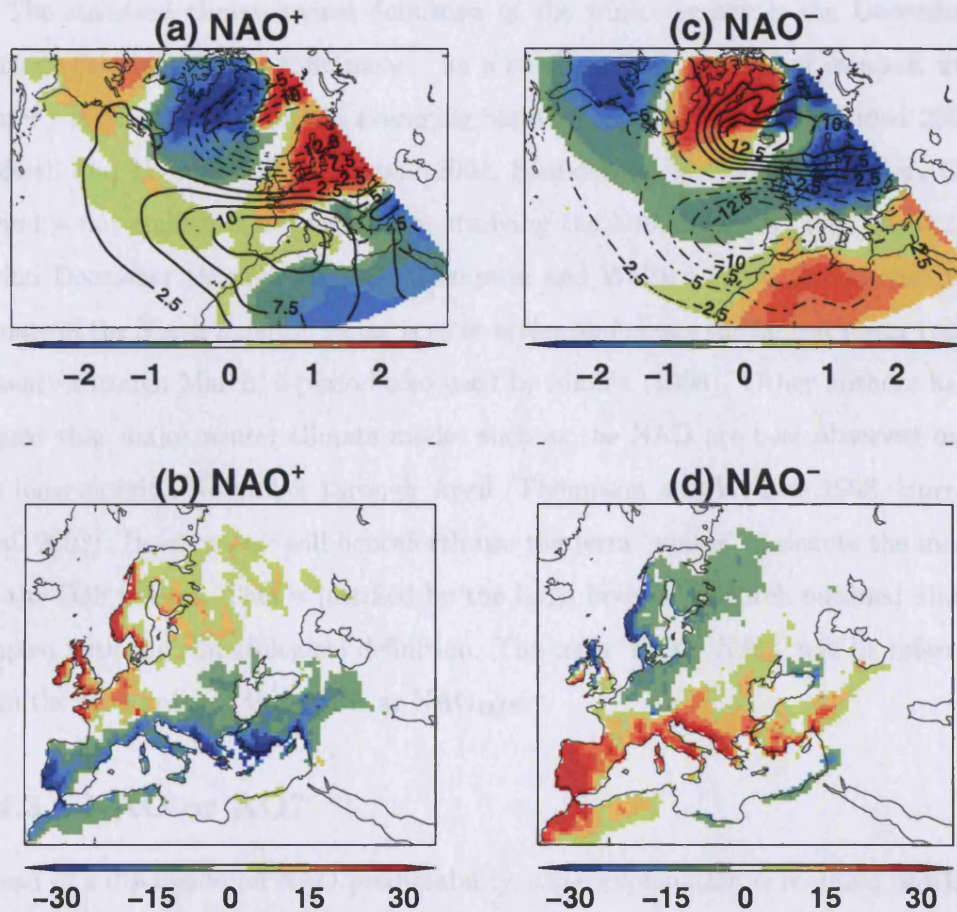


Figure 1.3: The temperature, mean sea-level pressure (MSLP) and precipitation effects of the positive (panels a and b) and negative (panels c and d) phases of the winter NAO. Panels a and c: contours represent DJF MSLP anomalies (contour interval is 2.5 hPa and negative contours are dashed), shading represents DJF 2 m NCEP/NCAR air temperature anomalies (shaded at 0.25, 0.5, 1, 2 and 3°C where blues are negative and reds are positive). Panels b and d: shading represents anomalies in DJF mean CRU 0.5° rainfall accumulation (shading interval is 5mm). Positive and negative NAO phases relate to mean of upper and lower NAO_{DJF} terciles 1951-2001. Anomalies are calculated with respect to the 1950-2001 mean. MSLP data are area weighted by the cosine of latitude.

global predictability (Mathieu et al. 2004). Third, it was shown in Section 1.3 that accurate NAO forecasts are of greatest benefit to the many potential end user groups during winter.

The standard climatological definition of the winter season is the December-January-February (DJF) time mean. As a result, a large volume of research into winter NH climate employs this averaging period (e.g., Branković and Palmer 2000, Rodwell and Folland 2002, Feldstein 2002, Saunders et al. 2003). However, this period is not employed universally. In studying the NAO, Hurrell (1995) used the period December through March. Thompson and Wallace (2000) found that the climate of the North Atlantic sector is most active and shows the largest linear trend January through March, a period also used by Shukla (1998). Other authors have argued that major winter climate modes such as the NAO are best observed over the longer period November through April (Thompson and Wallace 1998, Hurrell et al. 2003). However, we will henceforth use the term ‘winter’ to denote the mean for the DJF season. This is justified by the large body of research outlined above coupled with the climatological definition. The term ‘winter NAO’ will be referred to in the remainder of this thesis as NAO_{DJF} .

1.4.3 NAO or AO?

Ahead of a discussion on NAO predictability, some explanation is required of what exactly is meant by the term NAO. Recently, Wallace (2000) noted that the original definition of the NAO index by Walker and Bliss (1932) is more like the current definition of the Arctic Oscillation (AO) (Thompson and Wallace 1998) than the currently accepted definitions of the NAO. Figure 1.4(a and d) shows the patterns of 300 hPa geopotential height associated with the NAO and AO indices. The AO pattern is more zonally elongated and centred further to the west than the NAO. The meridional cross-sections (panels b, c, e and f) show that these differences are prominent throughout the depth of the atmosphere.

The NAO/AO debate concerns whether the AO is a true physical mode of variability or an artefact of performing statistical data reduction techniques on

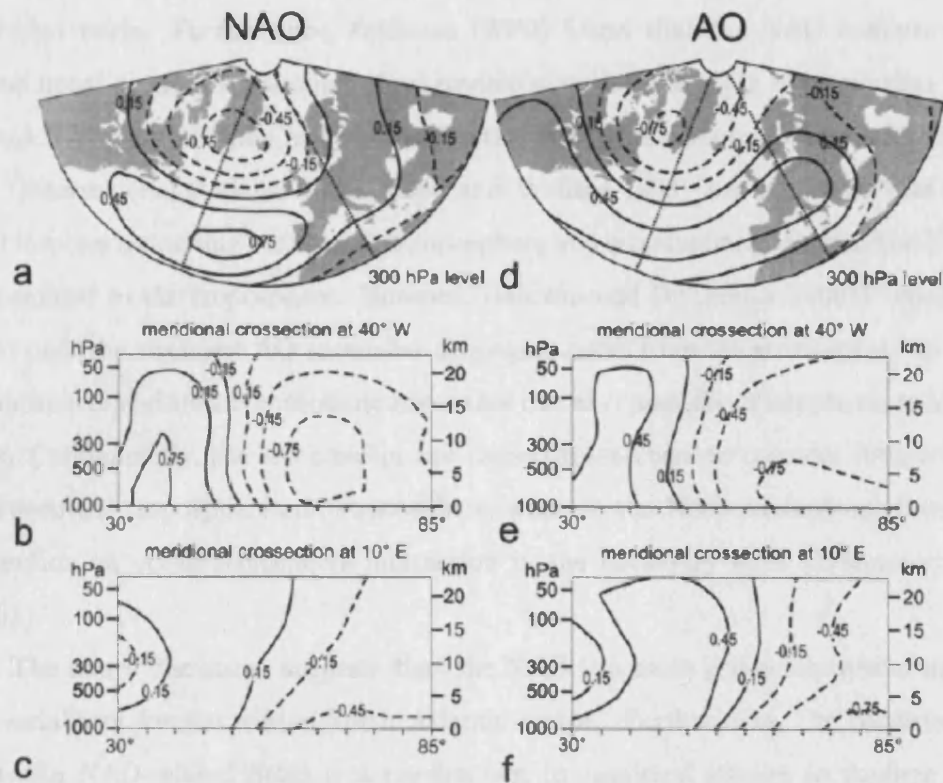


Figure 1.4: Correlations between monthly geopotential height anomalies and the NAO (left panels) and AO (right panels), respectively, for all months November to April 1958-97. Panels a and d show maps of the correlation at 300 hPa; panels b and e show meridional cross-section at 40°W; panels c and f show meridional cross-section at 10°E. Geopotential height data are from the NCEP/NCAR reanalysis, AO data are from Thompson and Wallace (1998) and NAO is based on index of Hurrell (1995). After Wanner et al. (2001).

hemispheric domains. Ambaum et al. (2001) showed that NH empirical orthogonal functions (EOFs) for streamfunction, zonal wind and temperature fields do not give dynamically consistent patterns. This means that it is not possible to decide which patterns are physically based and which result from the EOF technique. In contrast, these authors demonstrated that regional EOFs are physically consistent between different fields. Furthermore, Feldstein (2002) found that the NAO exhibits the same linear trend and variance characteristics as the AO. This suggests that the terms NAO and AO may in fact describe the same phenomenon (Wallace 2000).

Observational evidence (Thompson and Wallace 1998) has suggested that the AO involves a coupling between the troposphere and stratosphere, whereas the NAO is confined to the troposphere. However, Baldwin and Dunkerton (2001) reported that only the strongest AO anomalies propagate down from the stratosphere to the troposphere and that tropospheric anomalies can also precede stratospheric anomalies. Consequently, the AO concept has caused researchers to consider interaction between the troposphere and stratosphere, whereas the NAO concept has focused attention on ocean-atmosphere interaction in the boundary layer (Wanner et al. 2001).

The above discussion suggests that the NAO is a more physically useful mode of variability for the winter North Atlantic sector. Furthermore, the consistency between NAO-related fields is a requirement in empirical studies to confirm dynamical plausibility. The NAO also dominates in the lower troposphere over the North Atlantic, which is the main area of interest for this study. For these reasons, we will now focus on the NAO rather than the AO.

1.4.4 Boundary forcing of the NAO_{DJF}

There remains no clear consensus as to whether the NAO_{DJF} is a purely atmospheric phenomenon or whether it is influenced to some extent by external boundary forcing (Wanner et al. 2001). However, on interannual and decadal time scales, there is mounting evidence in observations (e.g., Radcliffe and Murray 1970, Czaja and Frankignoul 2002, Rodwell and Folland 2002, Saunders and Qian 2002) and from

GCM simulations (e.g., Rodwell et al. 1999, Sutton et al. 2001) to suggest that the NAO_{DJF} responds to changes at its lower boundary. This implies that the ‘white’ internal atmospheric power spectrum is ‘reddened’ by the lower-frequency changes at the underlying boundary. Hurrell and van Loon (1997) have shown that the NAO spectrum has become redder with time over the last century.

Local *in situ* forcing of the NAO_{DJF} is suggested to occur at the boundary with the North Atlantic Ocean. However, the specific region of the Atlantic that exhibits the strongest relationship with the NAO is a source of debate. Hurrell et al. (2004) proposed that the NAO_{DJF} responds primarily to variations in tropical or subtropical North Atlantic SST, whereas Radcliffe and Murray (1970) and Junge and Stephenson (2003) suggested SSTs southeast of Newfoundland are most important. Aside from local Atlantic SSTs, NAO_{DJF} teleconnections have been found with SSTs in other major ocean basins. Robertson (2000) found that SSTs in the southern Atlantic also influence the atmosphere in the northern extratropics. Hoerling et al. (2001) identified links between decadal and multi-decadal NAO_{DJF} variability and SSTs in the Indian and tropical Pacific oceans. Sutton and Hodson (2003) showed that interannual changes in the NAO_{DJF} may also be related to ENSO variability. Furthermore, ocean dynamics are also suggested to play a role in the decadal and multi-decadal relationships between the NAO_{DJF} and SSTs (Bjerknes 1964, Eden and Jung 2000).

Other boundary variables have been proposed as potential sources of NAO_{DJF} predictability. The first is the areal extent of snow cover for different regions of the NH. Cohen and Entekhabi (1999) proposed a link between autumn season Eurasian snow cover and the upcoming winter AO index. Increased snow cover is postulated to drive an amplification of orographically forced *in situ* stationary Rossby waves, which stimulate coupling between the troposphere and stratosphere. This coupling occurs in observations and in an atmospheric GCM when snow cover is allowed to vary freely but not when snow cover is fixed (Gong et al. 2002). Synoptically, increased snow cover is associated during autumn with a westward expansion of the Siberian high over Europe and the North Atlantic, which could influence the evolution of the NAO_{DJF} (Cohen and Entekhabi 1999).

More recently, Bojariu and Gimeno (2003) found that the link between Eurasian snow cover and NAO_{DJF} is more significant when snow anomalies from April to October are employed. They proposed that this reflects the important influence of snow albedo and hydrological feedbacks during spring and summer. Saunders et al. (2003) found that June-July mean NH snow cover is related significantly to the NAO_{DJF} . They proposed a dynamical link between snow cover and NAO_{DJF} through hemispheric-scale zonal gradients in subpolar air temperature, which are associated with persistent circulation anomalies and SST anomalies in the North Atlantic during subsequent months.

1.4.5 Lagged NAO_{DJF} predictability and skill

Davies et al. (1997) showed that only a maximum of 10-15% of the NAO_{DJF} variance can be explained by SST boundary forcing. While this is not a large proportion, any positive skill has value to end users because it represents an improvement over the intrinsically unpredictable day-to-day atmospheric variations. Therefore, attempts have been made at several research centres to produce seasonal forecasts for NAO_{DJF} .

Dynamical SCF models have shown some skill in reproducing the historical record of NAO_{DJF} variability. Rodwell et al. (1999) forced an atmospheric GCM ensemble with observed global SSTs and found a correlation of $r = 0.41$ between the observed and ensemble mean NAO_{DJF} indices (1947-1997). However, Bretherton and Battisti (2000) pointed out that caution must be used when interpreting results of this kind. They showed that the predictability of low frequency atmospheric variability increases with ensemble size. This is because the ensemble mean filters out the variability associated with internal atmospheric dynamics and leaves behind the coupled low-frequency signal. More reliable dynamical forecasts are produced using coupled GCMs (CGCMs) because the atmosphere and ocean evolve together in a physically consistent fashion. However, using the Météo France, UKMO and ECMWF CGCMs, the DEMETER ensemble mean correlation skill (henceforth r_s) for predicting NAO_{DJF} from November is only $r_s = -0.03$ for 1958-2001. The skill

improves to $r_s = 0.51$ when seven CGCMs are used over the shorter period 1980-2001 (Palmer et al. 2004), although most of the skill comes from the Météo France and CERFACS models.

Currently, the highest and most reliable skill for lagged NAO_{DJF} predictions comes from empirical studies. Recent studies have uncovered links between preceding late spring to autumn North Atlantic SST anomalies and the NAO_{DJF} . Rodwell and Folland (2002) highlighted a pattern of May North Atlantic SST that is related significantly to the upcoming winter mean 500 hPa geopotential height field, achieving a hindcast skill of $r_s = 0.45$ for the period 1948-1997. Saunders and Qian (2002) found that two modes of late summer/early autumn North Atlantic SST variability 1950-2001 were skilful in predicting a range of NAO_{DJF} indices with a hindcast correlation skill between $r_s = 0.47$ and $r_s = 0.63$. Relationships between lagged North Atlantic SSTs and NAO_{DJF} have also been highlighted by Drévilion et al. (2001), Czaja and Frankignoul (2002) and Cassou et al. (2004). These authors show that a horseshoe-like SST pattern in summer explains up to 16% of the early winter NAO variance 1958-1997.

Another lagged NAO_{DJF} predictor is the monthly mean areal extent of snow cover over different regions of the NH. Reliable satellite observations of snow cover are available only since 1972 (Robinson et al. 1993). Thus, the influence of anomalous snow cover on the winter atmospheric circulation and climate has received less attention than that of SST anomalies. Saito et al. (2001) highlighted the strong relationship between Eurasian snow cover in October and the upcoming winter AO index, which yields a correlation of $r = -0.55$ for 1972-1999. Saunders et al. (2003) used NH snow cover to predict a range of NAO_{DJF} indices and found that the June-July mean produced the most significant predictability with $r_s = 0.61$ for 1972-2001. A significant correlation of $r = -0.56$ is highlighted by Bojariu and Gimeno (2003) between April to October mean Eurasian snow cover and the NAO_{DJF} index 1973-1998. These studies suggest a recent significant link between NH climate in summer and that in the upcoming winter.

Figure 1.5 summarises the NAO_{DJF} skill currently available from dynamical and empirical models 1972-2002. Plotted are the observed NAO_{DJF} index and hind-

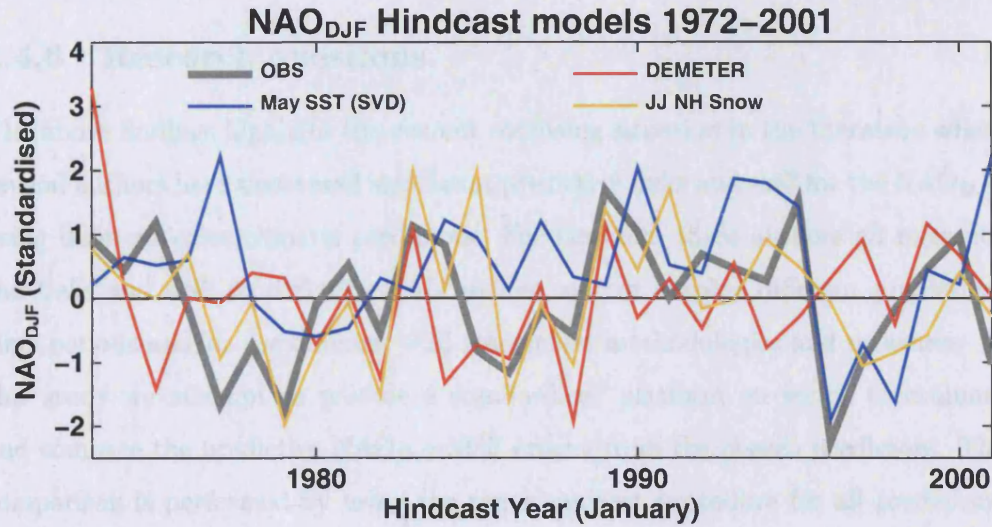


Figure 1.5: Cross-validated NAO_{DJF} hindcasts 1972–2001 from dynamical and empirical SCF models. Line colours denote: observations (CRU NAO_{DJF}) (grey), DEMETER ensemble mean (red), May SST (SVD) model (blue) and JJ NH Snow cover model (yellow). The NAO_{DJF} values are standardised with respect to the 1972–2001 mean.

casted NAO_{DJF} indices from the November to DJF DEMETER ensemble (Palmer et al. 2004), the May SST (SVD) model of Rodwell and Folland (2002) and the June–July (JJ) NH snow cover model of Saunders et al. (2003). Observations are from the CRU NAO_{DJF} index (Jones et al. 1997) and the Pearson correlations with the hindcasted NAO_{DJF} indices are $r = 0.17$, $r = 0.25$ and $r = 0.53$, for DEMETER, May SST (SVD) and JJ NH snow, respectively. The observed and hindcasted NAO_{DJF} time series are standardised to zero mean and unit variance for display purposes. In reality, both the DEMETER and statistical predictions underestimate the observed NAO_{DJF} variance. The hindcasted variances as percentages of the observed NAO_{DJF} variance range from 7% for DEMETER to 25% for the JJ NH snow cover model, which means that large departures from the mean are not well modelled. These results show that NAO_{DJF} hindcast skill is currently highest using empirical, rather than dynamical, SCF models. Therefore, we now focus on assessing lagged NAO_{DJF} predictability from empirical models.

1.4.6 Research questions

The above findings highlight the current confusing situation in the literature where several authors have uncovered significant predictive links and skill for the NAO_{DJF} using different prior climatic predictors. Furthermore, these authors all reference the links and skill to different NAO indices and/or employ different predictand time periods and/or use different skill assessment methodologies and measures. In this study we attempt to provide a standardised platform on which to evaluate and compare the predictive NAO_{DJF} skill arising from the chosen predictors. The comparison is performed by using the same hindcast procedure for all predictors. Hindcasts are made for three different NAO_{DJF} indices over three extended time periods out to 100 years and skill is assessed using two skill measures. Where possible, hindcast sensitivity to the choice of predictor observational data set is also examined by using two data sets for each analysis. These methods will allow us to provide answers to the following specific research questions, which are addressed in Chapter 3:

- (i) What is the most skilful lagged predictor of the NAO_{DJF} ?
- (ii) Are the predictive NAO_{DJF} relationships stationary when assessed over 100, 50 and 30 year periods?
- (iii) Does the most skilful period coincide with other periods of variability in either NAO_{DJF} or the predictors?
- (iv) What are the physical mechanisms that underpin the link between the NAO_{DJF} and the most skilful lagged predictor?

1.5 GCM representation of NAO links to prior Northern Hemisphere snow cover

As discussed in Section 1.4.5, summer NH snow cover was recently identified as a skilful lagged predictor of NAO_{DJF} (Saunders et al. 2003). Figure 1.6 shows the significant relationship between summer NH snow cover and the upcoming NAO_{DJF} . However, little is understood about the physics underpinning snow-climate interactions spanning warm to cold seasons. The mechanisms are probably different

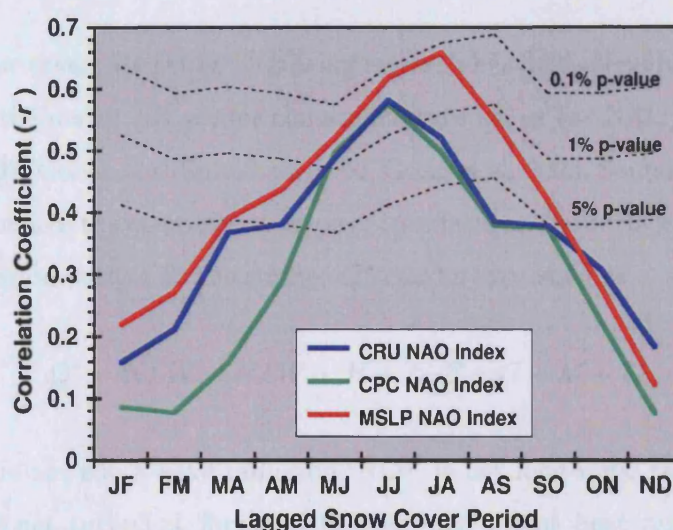


Figure 1.6: The correlation between lagged NH snow cover and three NAO_{DJF} indices 1972-2002 for bi-monthly snow cover periods ranging from JF (January-February) to ND (November-December). Negative correlation coefficients calculated with detrended time series are plotted. Dashed lines denote confidence levels corrected for serial correlation. After Saunders et al. (2003).

to those identified linking Eurasian snow cover and the winter climate, since the troposphere and stratosphere are uncoupled during summer (Saito et al. 2004). To date, GCMs have not been used to assess the summer snow/ NAO_{DJF} link. The aim of the second research topic in this thesis is to examine the output from a twentieth century coupled GCM integration and assess to what extent this link is captured. This will help to clarify the physical basis for the relationship.

1.5.1 The relationships between snow cover and climate

Contemporaneous relationship

Snow lying on the Earth's surface in large quantities has a profound effect on the atmosphere overlying the snow covered regions and on the whole climate system. Snow cover is the most variable NH land surface condition in both space and time (Cohen and Rind 1991). The major snow covered regions at all times of the year are located over northern and central Eurasia and northern North America. Blandford (1884) was one of the first to suggest a link between the evolution of NH climate

and prior snow cover. Recently, increasing research has focused on how lagged snow cover affects the major NH winter climate modes such as the NAO_{DJF} (Watanabe and Nitta 1998, Cohen and Entekhabi 1999, Gong et al. 2003, Saunders et al. 2003).

Snow influences the atmosphere through its effects on the surface energy budget. The net radiation budget at the surface Q^* can be expressed as

$$Q^* = NSW + NLW + H + L_V E + G + M + R, \quad (1.1)$$

where NSW is net short-wave radiation, NLW is net long-wave radiation, H and $L_V E$ are the net turbulent fluxes of sensible and latent heat and G is the net conductive ground heat flux, M is advective energy flux and R is the heat transfer due to precipitation. Over large spatial scales, the net contribution from G , M and R is small (Cline 1997). The dominant terms affecting the seasonal variability of Q^* on the spatial scales of seasonal predictability are the radiative (NSW and NLW) and turbulent (H and $L_V E$) fluxes. NSW at the surface is defined as

$$NSW = (K + k)(1 - \alpha), \quad (1.2)$$

where K is the incoming direct short-wave (solar) radiation, k is the downwelling diffuse short-wave radiation and α is the surface albedo. NLW at the surface is defined as

$$NLW = (L_D - L_U), \quad (1.3)$$

where L_D is downwelling long-wave (infrared) radiation and L_U is upwelling long-wave radiation.

Cohen and Rind (1991) outlined the main influences that snow cover can have on Q^* . First, the high reflectivity of fresh snow can increase α by 30-50%, thus reflecting a greater amount of incoming short-wave solar radiation before it is absorbed at the surface. Second, snow has a higher thermal emissivity than most other natural surfaces and therefore increases the amount of infrared radiation emitted to the atmosphere. Third, fresh snow acts as a thermal insulator because of its low thermal conductivity, preventing a positive flux of sensible heat from the ground

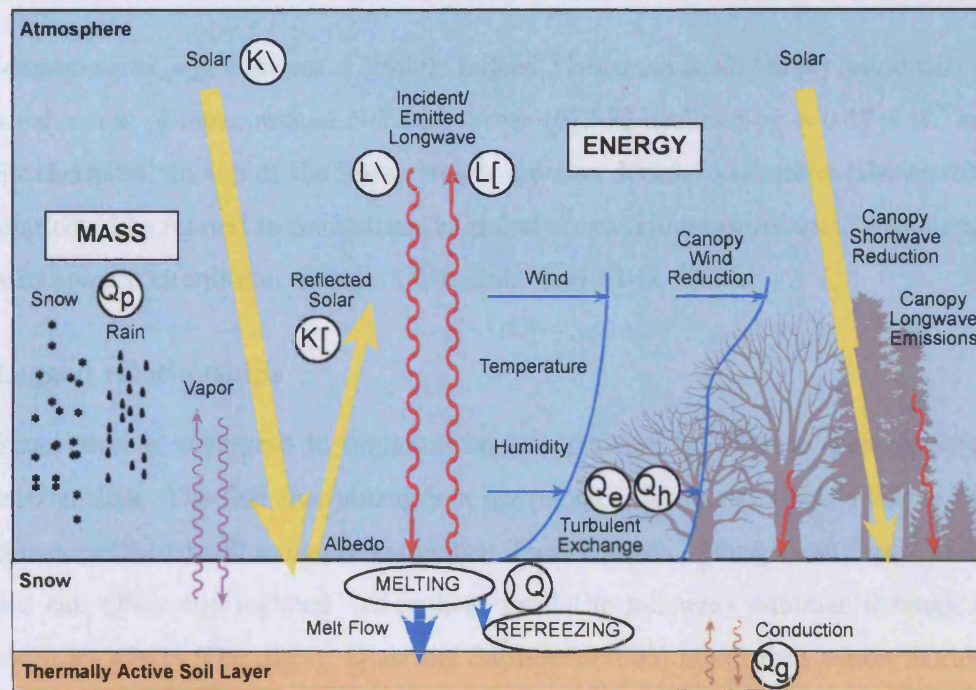


Figure 1.7: The energy budget above a snow covered surface. White circles denote transfers of energy: Q_e is latent heat flux through condensation or sublimation, Q_h is sensible heat flux through convection, Q_p is heat advected by rain falling on snow, “ Q ” is energy exchange within snow pack and snow melt, Q_g is soil heat flux through conduction. “ K ” is short-wave solar radiation, “ L ” is long-wave infrared radiation, “ \backslash ” denotes energy directed towards the surface and “[” denotes energy directed away from the surface. (Courtesy of D. Cline).

into the atmospheric boundary layer. Fourth, melting snow is a sink for latent heat. Fifth, Groisman et al. (1994) stated that snow cover suppresses convection. These effects are summarised in Figure 1.7.

Temporal and spatial variability

The spatial extent of snow cover varies in time more than any other climate variable (Cohen and Rind 1991). Annual variations in monthly mean NH snow cover swing between a maximum of $\sim 4.5 \times 10^7 \text{ km}^2$ in January to a minimum of $\sim 0.5 \times 10^7 \text{ km}^2$ in July (Robinson et al. 1993). Longer time scale trends are also known to be superimposed on the annual snow cover cycle. Much of the downward linear trend since the early 1970s is anticorrelated with increasing observed global mean

temperatures (e.g., Jones et al. 1999). Indeed, Groisman et al. (1994) found that the areal extent of mean annual NH snow cover 1973-92 declined by $\sim 0.17 \times 10^7$ km². Furthermore, on top of the linear trends are also decadal variations (Brown 2000) that could be related to oscillations in global ocean temperatures and/or associated atmospheric circulation changes (Watanabe and Nitta 1999).

Lagged relationships

Snow cover is suggested to impart a lagged influence on climate through several mechanisms. The first mechanism is a hydrological feedback, where melting snow causes persistent soil moisture anomalies. Over Eurasia, spring snow depth anomalies can affect the regional atmosphere until the following summer through soil moisture effects (Ose 1996). Qian and Saunders (2003) found that winter Eurasian snow cover also has a lagged teleconnected effect on upcoming summer UK and European temperatures. The delay in the atmospheric response results from modification to the surface radiation budget. Additional heat energy is required to warm the surface due to the increased specific heat capacity of the wet surface layers and to the latent heat energy required to evaporate the extra moisture (Walsh et al. 1985, Shinoda 2001). However, some debate exists as to the magnitude of the hydrological feedback. Shinoda (2001) and Robock et al. (2003) both conclude that the ‘memory’ of snow cover through soil moisture persists at the surface for a maximum of one month and at depth for a maximum of two months. This suggests that the direct association between melt season soil moisture anomalies and the subsequent climate over Eurasia is quite limited in duration. However, by modifying the surface energy budget over a large region such as Eurasia, even for one month, soil moisture anomalies could have a lagged impact on NH climate.

The second mechanism is a radiative feedback caused by the changing influence of snow cover on the surface energy budget during the annual cycle. Groisman et al. (1994) found that the NH net radiative balance at the top of the atmosphere has a distinct seasonal evolution that is linked to changes in the areal extent of snow cover. The dominant and competing terms in the energy budget are surface albedo

(α) and outgoing long-wave radiation (OLR). The radiative feedback is positive in autumn and early winter because decreased OLR outweighs increased α . Only from February onwards does the increased α dominate, which results in a net NH cooling.

The third mechanism is a dynamical feedback, where snow mass anomalies over a particular region (e.g., Eurasia) can significantly affect global atmospheric circulation patterns through hemispheric-scale wave trains (Kodera and Chiba 1989). Investigations into dynamical snow feedbacks have focused largely on the link between the Indian summer monsoon and prior Eurasian snow cover (e.g., Becker et al. 2001, Robock et al. 2003). Dynamical teleconnections have also been identified between the dominant mode of NH winter variability and prior Eurasian (e.g., Cohen and Entekhabi 1999, Watanabe and Nitta 1999, Cohen and Saito 2003) and NH (e.g., Saunders et al. 2003) snow cover. Figure 1.8 shows the detrended, standardised time series of NAO_{DJF} , October mean Eurasian snow cover and JJ mean NH snow cover. The Pearson correlation between NAO_{DJF} and the snow cover time series is $r = -0.37$ for October and $r = -0.58$ for JJ. These correlations suggest a physical link between lagged snow cover and NAO_{DJF} . However, 30 years of data is not sufficient to test rigorously the stationarity of these relationships.

The interactions between lagged snow cover and the winter climate modes may be a two-way process. Robock et al. (2003) showed that the evolution of Eurasian snow cover from winter to spring is negatively correlated with the preceding NAO_{DJF} . This suggests that reduced snowfall during positive NAO winters causes persistent warm temperatures during subsequent months, which maintain the reduced snow conditions. Cohen and Saito (2003) introduced similar arguments to explain the observed link between JJ NH snow cover and NAO_{DJF} . They stated that increased summer snow cover is indicative of a cool background climatic state, which persists until autumn when dynamical feedbacks occur between Siberian snow and the NAO_{DJF} . The coupled nature of the snow-atmosphere relationship provides an added challenge for GCM simulations.

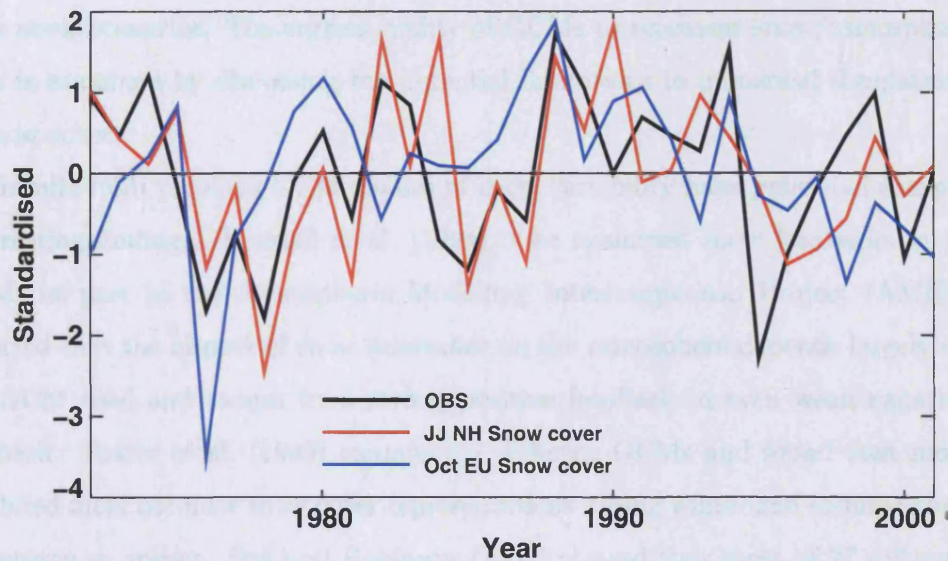


Figure 1.8: Time series of CRU NAO_{DJF} (black line), October mean Eurasian snow cover (blue) and JJ mean NH snow cover (red). Snow cover series are multiplied by -1 for display purposes. All time series are detrended and standardised with respect to the 1972-2001 mean.

1.5.2 GCM simulation of snow cover

Investigations into the interactions of snow cover and the climate system are restricted because observational data describing the areal extent of snow cover are spatially and temporally limited before the early 1970s when reliable satellite imagery became routinely available (Robinson et al. 1993). This limitation is even more acute during the summer season when *in situ* observations of snow cover at a particular location are inconsistent because of the reduced areal extent of snow cover (Brown 2000). Furthermore, empirically derived results are difficult to interpret because snow cover, air temperature and the atmospheric circulation are all interdependent. Cohen and Rind (1991) stated that observational studies are affected by the problem that if a winter with increased snow cover is colder than normal, is it colder because of the snow cover, or is there more snow cover because it is colder? In view of these limitations, the dynamical impact of snow cover on the atmosphere may be better assessed using GCM integrations forced with different

snow cover scenarios. The current ability of GCMs to represent snow/atmosphere links is examined by discussing the potential limitations in numerical simulations of snow cover.

Results from previous GCM studies of snow variability have presented sharply contrasting findings. Randall et al. (1994), who examined snow feedbacks in 14 GCMs as part of the Atmospheric Modelling Intercomparison Project (AMIP), reported that the impact of snow anomalies on the atmosphere depends largely on the GCM used and ranges from strong positive feedback to even weak negative feedback. Foster et al. (1996) examined 7 different GCMs and found that most exhibited more accurate snow cover representations during winter and summer than in autumn or spring. Frei and Robinson (1998) showed that most of 27 different AMIP GCMs reproduced an annual cycle of snow extent similar to observations but interannual variability was around one-half the observed magnitude. This was due to underestimation of snow cover during autumn and winter and overestimation during spring. This work also showed that snow cover in the AMIP GCMs, in contrast to observations, was unrelated to global SSTs. More recently, Frei et al. (2003) found that the second generation of AMIP GCMs (AMIP-2) produced a more realistic annual cycle but still underestimated the magnitude of interannual snow variability. They identified the major ongoing problem as regional-scale bias, particularly over the Eurasian continent.

The Coupled Model Intercomparison Project (CMIP) has evaluated systematically the accuracy of a suite of coupled GCMs (Covey et al. 2003). However, this project has yet to report on the representation of snow and related processes in the coupled GCMs. Specific comparisons between different land and snow surface modelling schemes have been performed, two examples being the Project for the Intercomparison of Land-surface Parameterisation Schemes (PILPS) (Slater et al. 2001) and snowMIP (Etchevers et al. 2002). However, these projects are performed ‘off-line’, which means that the snow models are assessed prior to being incorporated into a coupled GCM. Clearly, the interaction of the snow model with the other components of a coupled GCM presents a different set of problems than representing just snow processes.

Typically, snow accumulation and snow melt in GCMs are simulated by applying simple energy and mass balance accounting procedures (Foster et al. 1996). Recent improvements in GCMs have focused on adding vertical resolution and improved one-dimensional (vertical) physics. These algorithms frequently neglect or oversimplify important physical processes associated with subgrid-scale temporal and spatial variability of the snow covered area. These processes include wind blowing of snow, orographic influences, available insolation and vegetation distribution (Liston 2004). The lack of subgrid snow distribution representations in most GCMs has been identified as a deficiency in snow cover evolution and atmospheric interaction simulations (Slater et al. 2001). Snow cover in a GCM is typically specified to cover an entire grid cell homogeneously. Snow is seldom so uniformly distributed in nature, particularly at the snow margins, and this parameterisation has implications for the evolution of the surface energy budget. This is most important in grid cells over complex orography, where snow distribution is most variable in nature.

A further consideration when comparing results from observational and GCM snow studies is the difference between snow cover, snow mass and snow depth. While an area covered with deep snow is almost certainly uniformly covered, differences arise when the amount of snow is small (Frei and Robinson 1998). Furthermore, a deep snow layer has different effects compared to a thin layer on the sensible and latent heating terms of the surface energy budget. The standard output in GCMs is snow depth or mass because this is the most important variable in relation to the other climatic parameters within the model. However, the observational data describe the areal extent of snow cover and so the two are not readily comparable in all cases. Therefore, in order to compare the observational and GCM responses to snow cover, an index representing GCM snow cover is required.

1.5.3 GCM simulation of snow/atmosphere links

As with previous observational studies, the majority of GCM simulations involving snow cover have concentrated on Eurasian snow links to the Indian summer monsoon (e.g., Vernekar et al. 1995, Liu et al. 2004). However, recent investigations

have focused on the lagged influence of NH regional and hemispheric snow cover on the upcoming winter atmosphere, including over the Atlantic sector. Gong et al. (2003) used a 20-member atmospheric GCM ensemble to examine the relationship proposed by Cohen and Entekhabi (1999) linking autumn Siberian snow cover and upcoming winter NH climate. The GCM was forced with realistic high and low snow cover extents over Siberia. Dynamically, the ensemble mean for high Siberian snow showed increased local anomalous upward wave activity associated with a weakened stratospheric vortex and a negative AO anomaly. The spatial pattern of the GCM response correlated strongly with observations but was only one-third of the magnitude. This weakened response may result from errors in the GCM formulation or from artificial damping by the ensemble averaging process. To our knowledge, this thesis is the first study to employ data from a coupled GCM to examine teleconnected links between summer snow cover and the atmosphere of the North Atlantic sector.

The above discussion, and that in Section 1.4.5, highlights the debate concerning the exact nature, magnitude and timing of the NH winter atmospheric response to lagged snow cover. This subject has received less attention than the role of lagged SST anomalies. Previous observational studies imply but do not conclusively prove a causal relationship between lagged snow cover and winter climate (Gong et al. 2003). Furthermore, many of these studies achieve results whose statistical significance is marginal (Fasullo 2004). The stationarity of these results has also yet to be determined. Previous studies using atmospheric GCMs to simulate snow/atmosphere links have shown promising results. This suggests that using a coupled ocean-atmosphere GCM may provide increased information about NAO_{DJF} predictability from lagged summer snow cover. These factors provide the motivation for the second research topic in this thesis.

We will perform a verification of the dynamical teleconnection proposed by Saunders et al. (2003), linking NH summer snow cover and NAO_{DJF} . Our methodology differs from that of other GCM snow studies (e.g., Gong et al. 2003) because data are taken from a single coupled GCM (CGCM) integration. The CGCM is run with observed twentieth century radiative forcing and with all other parameters allowed

to evolve freely. Using this approach, a snow/NAO_{DJF} link will be observed only if the CGCM formulation is sufficiently realistic and exhibits a strong signal-to-noise ratio. For this reason, we first perform a verification of the accuracy of spatial and temporal snow cover variability within the CGCM. This provides the first detailed verification of twentieth century snow cover variability in a CGCM. This study forms a link between previous observational and atmospheric GCM studies and potential future CGCM perturbation experiments. Therefore, this research could be useful to dynamical modellers developing future generations of CGCMs. Furthermore, this analysis will allow the stationarity of the observational findings to be tested in a realistic CGCM simulation.

1.5.4 Research questions

Dynamical model simulations of lagged snow cover relationships with the NAO_{DJF} are shown above to be limited both in number and, potentially, by the accuracy of individual GCMs. The performance of coupled GCMs in reproducing observed patterns of spatial and temporal snow cover variability has not been determined rigorously. In Chapter 4, we attempt to validate the accuracy of a coupled GCM for snow/atmosphere investigation by answering the following questions:

- (i) Can CGCM snow depth data be employed to create an index of monthly mean snow cover?
- (ii) How well does a coupled GCM simulate the observed spatial and temporal variability in seasonal snow cover when forced with observed radiative forcing 1972-2002?
- (iii) Does the coupled GCM simulate realistically the observed contemporaneous *in situ* links between seasonal snow cover and the atmosphere 1972-2002?
- (iv) Does coupling a GCM to the ocean improve its representation of snow cover variability?

The findings from Chapter 4 will show whether a coupled GCM is accurate at representing snow cover and associated atmospheric variability. In Chapter 5 we test the capability of the coupled GCM to represent NAO_{DJF} predictability associated with summer NH snow cover. The following specific questions are addressed:

- (i) Does a coupled GCM capture observed temporal NAO_{DJF} variability during the twentieth century?
- (ii) Does a coupled GCM represent accurately the observed lagged and contemporaneous links between snow cover and the NAO_{DJF} 1972-2002?
- (iii) Is the physical mechanism linking summer snow cover and the NAO_{DJF} the same in the coupled GCM as in observations?
- (iv) Is the link between summer snow cover and the NAO_{DJF} stationary during the twentieth century?
- (v) What are the main limitations in using a coupled GCM for snow/atmosphere modelling?

1.6 Summary

The contrasting methodologies employed in the seasonal forecasting of winter North Atlantic climate are described. Empirical models currently offer higher and more reliable seasonal forecast skill. However, increasing research is directed towards improving dynamical seasonal forecast systems. This study will focus on the seasonal-to-interannual predictability of the winter North Atlantic Oscillation (NAO) available from empirical models. The impacts of the winter NAO are clear, as are the benefits of accurate and timely seasonal NAO forecasts. Particular beneficiaries will be energy providers, water resource managers, farmers and the insurance industry.

The research divides into two main topics. First, several previous empirical modelling studies have identified predictive links and skill for the winter NAO. This work will clarify what is the best lagged predictor for the winter NAO. Second, summer snow cover has been identified in observations as a skilful lagged winter NAO predictor. This study uses output from a coupled GCM to validate this link and assess its stationarity during the twentieth century.

Chapter 2

Data Selection

2.1 Introduction

Several hundred years of meteorological observations have led to the accumulation of an enormous body of climate data. The aim of this Chapter is to identify which of the many available climate data sets best suit the data requirements of this thesis. These requirements are for the data to have little bias and the best available temporal and spatial coverage for the parameters of interest. The available historical data sets are examined and a review of the methods involved in their formulation is presented. This will allow an informed selection to be made of the most appropriate data sets for use in this study.

Meteorological and oceanographic data are taken from two main sources. First, direct measurements are made at observing locations on the ground or at sea. These data provide the longest records, extending (occasionally) back several hundred years, for example, the Central England Temperature index (Manley 1974). However, there is considerable bias in the density of the observation network towards the industrially developed regions of the world such as Europe and North America. Second, remote sensing satellites in orbit around the Earth detect radiances emitted from the surface. Satellite observations provide global coverage but records do not extend further back than the early 1970s. Long data records are essential for empirical SCF applications to determine the stability of predictive relationships. However, remotely sensed data play an increasingly important role in the observation and reconstruction of mean sea-level pressure (MSLP), air temperature, sea surface temperature (SST) and snow cover data.

The required climate parameters for this thesis are the predictands (what we predict) and the predictors (what we use to predict). The main predictand is the winter NAO (NAO_{DJF}). Section 2.2 examines available global MSLP data, which are employed to calculate the NAO_{DJF} indices. The methods used to compute the various NAO_{DJF} indices are discussed in Section 2.3. The chosen predictors comprise lower boundary variables including near-surface air temperatures, SSTs and snow cover extent. We review data for, air temperature in Section 2.4, SST in Section 2.5 and snow cover in Section 2.6. The Chapter concludes with a summary of the data sets selected for use in this thesis.

2.2 MSLP data

Observations of MSLP at land stations and aboard ships have been performed routinely since the mid-1800s. Interpolating those data into MSLP fields with complete global coverage has been more problematic. Reliable satellite and reanalysis products have increased the accuracy of complete global MSLP data for the period since 1980. However, the reliability of global reconstructed MSLP data sets based on surface observations is limited before around 1950 because of data sparsity in certain regions. Furthermore, these data sets are commonly available over the ocean only.

Measurement

Barometry is concerned with measuring static pressure, which is the force per unit area exerted by the atmosphere against any surface in the absence of air motion (Brock and Richardson 2001). Consequently, the instrument used to measure static air pressure is the barometer. The SI unit of atmospheric pressure is the Pascal (Pa), although in meteorology use of the millibar (mb) is still common. As a result, atmospheric pressures are typically given in hundreds of Pascals (hectopascals, hPa), where $1 \text{ hPa} = 1 \text{ mb}$.

Barometers used for meteorological observations are typically either mercury or aneroid. Mercury barometers work by the weight of the atmosphere balancing against a column of mercury in a long glass tube. The movements of the top of

the mercury column are calibrated to represent a certain change in atmospheric pressure. Mercury barometers are known to be stable in time, which is important for climatological use. However, they are very sensitive to motion and difficult to automate, making them unsuitable for field use.

Aneroid barometers are fluidless and work via changes in atmospheric pressure causing a metallic diaphragm to deflect inside an evacuated chamber. The metallic diaphragm sensor is a circular capsule made from an elastic metal. Increasing atmospheric pressure causes greater deflections and these are calibrated empirically at assembly. The smaller design and metallic construction make aneroid barometers more suitable for automated weather stations and in the field.

Errors

Each barometer design has associated errors and inhomogeneities that must be considered. Mercury barometers suffer mainly from dynamical wind pressure superimposed on the static pressure, which can result in errors of several hPa (Brock and Richardson 2001). These errors are difficult to trace due to the time-varying nature of dynamical wind pressure. Contamination in either the mercury or the vacuum at the top of the glass tube can also lead to inconsistency in the readings. Finally, the storage and maintenance of a mercury barometer is critical and it must be kept vertical and fixed.

The main source of error in aneroid barometers is from the nonlinear response of the barometer to changes in ambient air temperature. A temperature correction is incorporated at assembly, however, due to the many sources of temperature errors, there is always some residual temperature error (Brock and Richardson 2001). A hysteresis effect can cause significant problems through defects or irregularities in the diaphragm shape or material. Finally, drift caused by diaphragm creep can only be suitably corrected by periodic checking and offset correction.

Data set	Time period	Resolution	Coverage	Reference
ERA-40 ^a	1958 to present	$\sim 2.5^\circ \times 2.5^\circ$	Global	ECMWF (2003)
ERSLP (Recon) ^b	1854 to 1997	$2^\circ \times 2^\circ$	Ocean	Smith and Reynolds (2004b)
GMSLP2.1f (BP97) ^c	1871 to 1994	$5^\circ \times 5^\circ$	Global	Basnett and Parker (1997)
Kaplan (KEA) ^d	1854 to 1992	$4^\circ \times 4^\circ$	Ocean	Kaplan et al. (2000)
NCEP/NCAR ^e	1948 to present	$2.5^\circ \times 2.5^\circ$	Global	Kalnay et al. (1996)

^a ECMWF 40 year reanalysis.

^b Extended Reconstructed Sea Level Pressure.

^c UKMO Hadley Centre Global Mean Sea Level Pressure (version 2.1f).

^d Kaplan Reduced Space Optimal Interpolation.

^e NCEP/NCAR reanalysis.

Table 2.1: Global analyses of sea level pressure data. Data set titles in brackets are those given by Smith and Reynolds (2004b) and allow comparison with Figure 2.1.

Global gridded data sets

Five global reconstructed MSLP data sets are listed in Table 2.1. The data set titles in brackets are the titles given by Smith and Reynolds (2004b) and are used in the following to allow a comparison with Figure 2.1. Observational data over the ocean in Recon and KEA are taken from the Comprehensive Ocean-Atmosphere Data Set (COADS) (Woodruff et al. 1998). COADS data are from a network of buoys and ship observations. GMSLP2.1f is the only MSLP data set offering complete global coverage over both land and ocean for the entire twentieth century. Oceanic observations are taken from COADS and the UKMO Marine Data Bank, while land observations are taken from 700 historical station records and several oceanic or ‘near global’ gridded MSLP data sets. The reader is referred to the references in Table 2.1 for details of the exact methods used in the computation of each data set.

The spatial correlation between four of the MSLP data sets (over the ocean only) listed in Table 2.1 is shown in Figure 2.1. National Centers for Environmental Prediction (NCEP)/National Center for Atmospheric Research (NCAR) reanalysis data are plotted from 1980 onwards to coincide with the satellite era. The lowest correlations between the data sets are observed prior to 1950, when sampling is more sparse. The correlation between Recon and GMSLP2.1f is lowest in all time

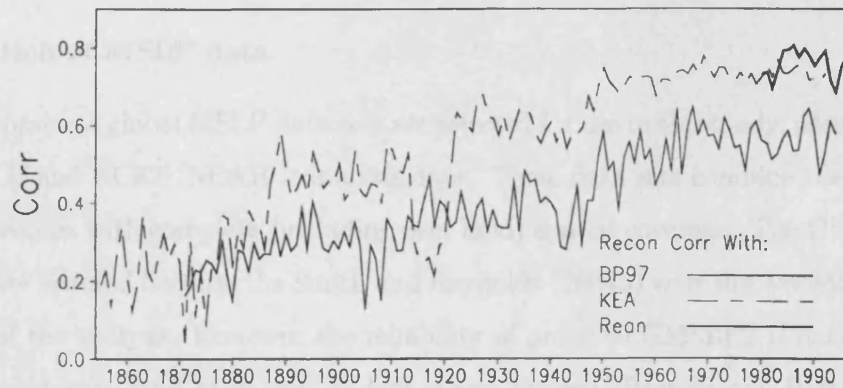


Figure 2.1: Comparison of global reconstructed MSLP data sets. Lines show annual spatial Pearson correlations (60°S - 60°N) between the Smith and Reynolds (2004b) data ("Recon"), Basnett and Parker (1997) (thin solid line, "BP97"), Kaplan et al. (2000) (dashed line, "KEA") and NCEP/NCAR reanalysis (thick solid line, "Rean"). After Smith and Reynolds (2004b).

periods, which is attributed to the reduced smoothing and greater noise content in the GMSLP2.1f data (Smith and Reynolds 2004b). NCEP/NCAR reanalysis data assimilate many different types of data from *in situ* observations and satellites and are considered to be more accurate than the historical reconstructions since the 1980s (Smith and Reynolds 2004b).

Smith and Reynolds (2004b) outlined the main sources of error in global reconstructed MSLP data:

- (i) Data quality problems, particularly over land. These include discontinuities between data sets related to the corrections applied to high-altitude stations.
- (ii) Quality control procedures were applied to the observational data to remove noise and gaps. A subjective threshold of 5 standard deviations was used to classify outliers, and the results showed some sensitivity to different thresholds.
- (iii) Trends caused by changes in data sampling. One interesting issue is that better weather forecasts allowed ships to avoid severe storms and, inadvertently, introduced a bias to MSLP sampling over the ocean. The removal of these artificial trends from the data may be sensitive to the type of corrections applied.
- (iv) The Earth's apparent gravity changes slightly from the equator to the Pole, creating the need for minor corrections ($\sim \pm 2$ hPa).

Selection of MSLP data

Two complete global MSLP data sets are selected for use in this study, namely, GMSLP2.1f and NCEP/NCAR reanalysis data. These data sets combine the need for long records with complete (including over land) spatial coverage. The GMSLP2.1f data are selected because the Smith and Reynolds (2004b) were not available at the time of the analysis. However, the reliability of pre-1950 GMSLP2.1f data may be poor over some land areas and in data sparse regions (Basnett and Parker 1997). Furthermore, there is greater noise in the data than in the Recon data (Smith and Reynolds 2004b). The results from these periods and regions must therefore be interpreted with some caution.

2.3 NAO indices

There is no single way to ‘define’ the NAO (Hurrell et al. 2003), and consequently there are many different methods used to compute NAO indices. Walker and Bliss (1932) were among the first to construct an NAO index, which related temperature and MSLP anomalies from European and North Atlantic observing stations. The NAO indices referenced commonly in the literature employ different data sets and statistical methods in their computation but most are derived from MSLP.

Station based NAO indices

The simplest form of NAO index is calculated as the difference in MSLP at stations located within the opposing centres of the NAO dipole pattern (see Figure 1.4a). The MSLP data are typically standardised to reduce the dominance of the higher latitudes caused by their greater variability. The northerly centre is termed the ‘Icelandic Low’ and the MSLP record at Reykjavik is used. The southerly centre is termed the ‘Azores High’ but this centre has a larger spatial extent than the Icelandic Low and, consequently, there is no universally agreed station to represent it. Hurrell (1995) used MSLP from Ponta Delgada in the Azores (henceforth Hurrell NAO index), while Jones et al. (1997) used MSLP from Gibraltar (henceforth CRU

NAO index). These NAO indices exhibit certain differences, which represent the temporal and spatial variability in the NAO dipole pattern. This variability results from the seasonal migration in the relative positions of the Icelandic Low and Azores High (Hurrell et al. 2003).

One advantage of using a station based approach is that long and accurate records of MSLP are available (e.g., from 1824 for Iceland). Furthermore, proxy data have also been used to reconstruct an NAO index back to 1500 (Luterbacher et al. 2002). The disadvantage of using station data is that the temporal and spatial variability of the NAO dipole are nonstationary (Barnston and Livezey 1987). Therefore, an index that is not tied geographically to two specific stations may be more representative of NAO variability.

Spatially varying NAO indices

One method used for overcoming the problem of spatial variability in the NAO dipole is principal components analysis (PCA). In this approach, the NAO is computed by projecting the leading eigenvector (EOF) of the North Atlantic MSLP cross-covariance matrix back onto the original gridpoint values of MSLP (Hurrell et al. 2003). Wallace (2000) states that the early NAO index of Walker and Bliss (1932) can be considered an iterative approximation to PCA.

PCA is considered to produce a better representation to the NAO index than a simple station based approach. This is because the EOF weights take into account seasonal migration in the two NAO centres of action. Furthermore, the EOF weights the MSLP time variability over regions where there may be no stations located. Hurrell (1995) was one of the first to employ an NAO index based on the PCA of North Atlantic sector winter MSLP fields (henceforth MSLP PC1 NAO index).

There are certain disadvantages to the PCA approach. First, the data used in the PCA are typically grids of reconstructed or reanalysis MSLP data sets. These data may not be reliable over land or over the ocean away from major shipping lanes (Smith and Reynolds 2004b). Paradoxically, the only way to ensure that reconstructed data are accurate is through comparison with observations and, by

definition, these data sets are least accurate where observational data are sparse.

Second, while PCA captures seasonal and interannual migration in the NAO pattern, it tends to smooth out spatial variations occurring over longer time periods. Several studies have highlighted that the mean position of the NAO dipole is nonstationary on decadal and longer time scales (e.g., Jung and Hilmer 2001). Therefore, the EOF pattern is sensitive to the time period being assessed. Hurrell (1995) computed the EOF using winter MSLP data 1899-2001, while other authors have employed shorter time periods (e.g., Saunders and Qian 2002). This means that the PCA-based NAO indices computed in different studies may not be comparable.

Selection of NAO indices

In order to provide a rigorous assessment of winter NAO prediction skill, three NAO indices are selected for this study. They are the CRU, Hurrell and Hurrell MSLP PC1 indices, respectively. The selection of two station based indices and one spatially varying index provides a set of predictands that cover a large area of the North Atlantic. The chosen predictors must therefore influence the atmosphere on sufficiently large spatial scales to provide predictability for all three NAO indices.

2.4 Near-surface air temperature

Historical records of temperature are the longest of all meteorological records. Accurate and reliable monthly mean temperature data are available for the UK dating back to 1659 (Manley 1974). Compared to other parameters such as precipitation, temperature varies more slowly and predictably in space and time. This means that fewer stations are required to achieve a representative observing network of the temperature conditions over a particular region. However, orographic effects and proximity to water bodies can cause air temperatures to vary much more rapidly over short distances (e.g., van den Dool and Nap 1981). Satellite retrievals since the early 1980s have made uniform global air temperature measurements available

for the first time, but debate exists as to the reliability of some of these data (e.g., Spencer and Christy 1990, Fu et al. 2004).

Measurement

In situ

Air temperature is one of the most fundamental of all meteorological parameters. The most common and simple method employed for temperature measurement is the liquid-in-glass thermometer. Several alternative instruments for measuring *in situ* air temperature are reviewed in Brock and Richardson (2001). However, the majority of global climatological near-surface temperature observations are still made using manual thermometers.

Satellite

Remote sensing of air temperature is performed using infrared and microwave sounding instruments. Satellite-borne sensors detect radiances emitted from an atmospheric layer, which are linearised based on an *a priori* estimate of the temperature profile to yield temperature data. Near-surface air temperatures are generally estimated using an empirical relationship with surface (or skin) temperature (see also Section 2.5). Infrared instruments such as the high resolution infrared radiation sounder (HIRS) have good horizontal and vertical resolution but suffer from the attenuation of radiances by clouds at infrared wavelengths (e.g., 4.3 μm). Microwave sounding units (MSU) have the advantage that the transparency of clouds is much increased at microwave frequencies (e.g., 50.31 GHz). However, their limitation is that the vertical and horizontal resolution is ~ 10 times coarser than HIRS. Research has attempted to exploit the advantages of HIRS and MSU through a combination of the data from both instruments (Goldberg and McMillin 1999). Satellite observations are therefore becoming more useful for retrieving areal estimates of temperature, but thermometer measurements remain the only source of observational data in many regions of the world.

Errors

Errors in the measurement of air temperature in excess of 2 to 3°C are not uncommon in many networks (Brock and Richardson 2001), which are usually acceptable for the general public. However, numerical models at all scales of motion are greatly affected by errors even as large as 1°C due to their sensitivity to initial conditions (e.g., Crook 1996). Temperature data errors are often caused by changes in thermometer design. However, other sources of inhomogeneity and bias also exist. The most common problem is observer error. This may occur either during measurement or in the transcription of records. Comprehensive error detection requires comparisons with either nearby station data, station metadata, documentary records or even reference to the original registers. Since this is time consuming and not practical for large data sets, automated procedures have been developed. These are based on the identification of data points outside a predefined confidence interval for the statistical distribution representative of the temperatures at the point in question.

Errors in remotely sensed near-surface temperatures stem from two sources. First, the retrieval process is based on converting nonlinear emitted radiances to linear estimates of atmospheric temperature. This is achieved using weighting functions that are chosen to minimise the thickness of the atmospheric layer they represent. However, the weighting functions are imperfect and radiances detected from a given layer also contain information from neighbouring layers. These errors are removed through empirical estimates of the temperature profile. Second, near-surface layer radiances cannot be detected precisely by either infrared or MSU instruments. Therefore, near-surface air temperatures are typically inferred from surface (skin) radiances. The relationship between surface and near-surface air temperature can only be verified through *in situ* observations, which means that errors are more likely over data sparse regions.

Global gridded data sets

The irregular distribution of available station data requires that some form of gridding is necessary, in order for large-scale temperature analyses (e.g., hemispheric averages, principal component patterns etc) not to be biased (Jones and Moberg 2003). One common method of interpolating station temperature data onto a regular grid is the 'climate anomaly' approach (e.g., Jones 1994). This method reduces all the station temperature data to monthly anomalies with respect to a common period, such as 1961-90. Gridbox anomaly values are then produced by a simple averaging of the individual station anomaly values within each grid box. The advantage of this method is that localised differences between stations (e.g., through elevation effects) are minimised. The major disadvantage is that stations must have enough years with data within the common period in order to be used. Clearly, this method is applicable to other climatic parameters that are measured systematically at a sufficiently dense network of surface stations (e.g, wind speed and precipitation).

Several independent global gridded temperature data sets have been produced (Table 2.2). Complete global coverage is offered in the reanalysis data from NCEP/-NCAR (Kalnay et al. 1996) and the ECMWF 40-year reanalysis (ECMWF 2003). In the reanalysis system, station and (post 1980) satellite observations are assimilated into a coupled GCM integration. The relative weights given to the observational data and the model output are determined by the quality and quantity of observations for a certain area at a given time. Therefore, in the pre-satellite era, model-related biases may exist over data sparse regions. The ERA-40 data have certain advantages over NCEP/NCAR, such as a higher top stratospheric level (1 hPa compared to 10 hPa in NCEP/NCAR). However, ERA-40 is still relatively new and has not been verified as extensively as the NCEP/NCAR data.

There are also global gridded temperature data sets without complete coverage. CRUTEM2 (Jones and Moberg 2003) has data over land only but offers historical records over all five continents dating back to 1856. The data contain a high percentage of missing grid cells, particularly in the Southern Hemisphere. Cover-

Data set	Time period	Resolution	Coverage	Reference
CRUTEM2 ^a	1856 to present	5° × 5°	Land only	Jones and Moberg (2003)
ERA-40 ^b	1958 to present	~ 2.5° × 2.5°	Global	ECMWF (2003)
COADS ^c	1871 to 1994	5° × 5°	Ocean only	Woodruff et al. (1998)
NCEP/NCAR ^d	1948 to present	2.5° × 2.5°	Global	Kalnay et al. (1996)

^a Climatic Research Unit, University of East Anglia.

^b ECMWF 40 year reanalysis.

^c Comprehensive Ocean Atmosphere Data Set.

^d NCEP/NCAR reanalysis.

Table 2.2: Global analyses of 2 m air temperature data.

age is good throughout the record over Europe and North America but is limited over eastern Eurasia prior to the 1930s. COADS data are based on ship and buoy observations (Woodruff et al. 1998) and are therefore unavailable over land.

Choice of air temperature data

The requirements for air temperature data in this thesis concern their temporal and spatial coverage. The data must span a period of at least 100 years to allow predictability studies with the long NAO records. Furthermore, a significant percentage of the NH land must be covered to allow teleconnected climatic relationships to be assessed. Two data sets satisfy one or more of these criteria, namely, CRUTEM2 and the NCEP/NCAR reanalysis data. A comparison of NCEP/NCAR and CRUTEM2 data for NH regions poleward of 35° is shown in Figure 2.2. The large number of missing data in CRUTEM2 is apparent, but over the common areas with data the correspondence with NCEP/NCAR is high. Inconsistencies between the data sets probably arise through differences in their spatial resolutions and the CGCM output in the NCEP/NCAR data.

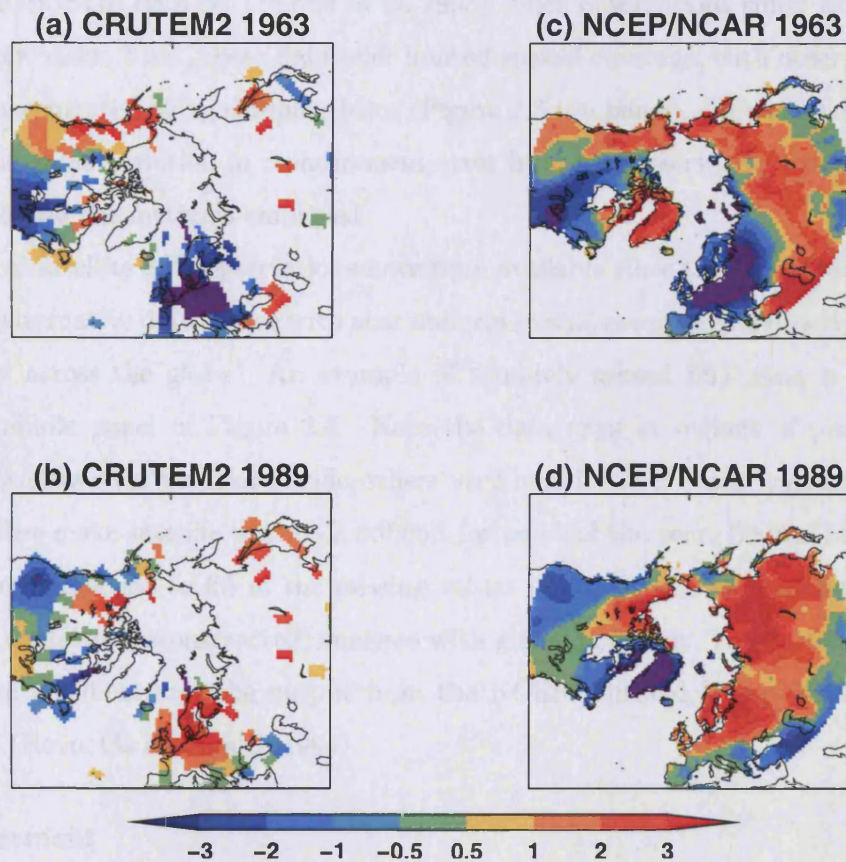


Figure 2.2: DJF mean 2 m air temperature anomalies from CRUTEM2 (a and b) and NCEP/NCAR reanalysis (c and d) data. Data are for 1963 (top panels) and 1989 (bottom panels), which correspond, respectively, to extreme negative and positive NAO phases. Missing data in CRUTEM2 are white space. Colour levels are at 0.5, 1, 2, 3 and $> 3^{\circ}\text{C}$, where blues indicate negative and reds indicate positive temperature anomalies.

2.5 Sea surface temperature

Historical records of SST extend back to the mid 1800s. In the pre-satellite era, SSTs were measured *in situ* by ships and buoys. The primary *in situ* observational data sources are the COADS (Woodruff et al. 1998) and the UKMO Historical SST (MOHSST6) data set (Parker et al. 1995). Such observations suffer from two main drawbacks. First, these data offer limited spatial coverage, with observations being concentrated along shipping lanes (Figure 2.3 top panel). Second, there may be considerable variation in measurement error from one observation to the next, depending on the methods employed.

Global satellite SST observations have been available since the early 1980s, and offer an alternative data source with near uniform spatial coverage and observational accuracy across the globe. An example of remotely sensed SST data is shown in the middle panel of Figure 2.3. Note the data gaps in regions of persistent cloud cover over the tropical Pacific, where very high levels of water vapour in the atmosphere make satellite retrievals difficult for much of the year. Several methods have been employed to fill in the missing values in the *in situ* and satellite data to yield complete ‘reconstructed’ analyses with global coverage. The bottom panel of Figure 2.3 illustrates the output from the NCEP Optimal Interpolation (OI) analysis (Reynolds and Smith 1994).

Measurement

In situ

Historically, SST has been referenced to a bulk near-surface ocean temperature obtained by tossing a bucket over the side of a ship. Since the design and insulation of buckets has changed with time, corrections must be applied (Folland and Parker 1995). During World War II, engine room intake-water thermometers superseded bucket measurements. These temperatures are affected by the depth of the ship’s intake, the loading of the ship, the configuration of the engine room and the point where the measurement is taken. Such differences contribute to noise in the SST measurements. However, biases also arise because the engine room heat more than



Figure 2.3: A comparison of global SST coverage for a week in December 1999 based on: *In situ* observations (top), satellite observations (middle) and the NCEP OI analysis model (bottom). The colour bar indicates the SST anomaly in °C. (Courtesy of R. W. Reynolds).

offsets any cold bias from the depth of the intake. Overall, the difference between engine intake and bucket temperature is typically 0.3°C (see also Trenberth et al. (1992) and Rayner et al. (2003)).

Satellite

At no stage in the observational record have *in situ* SST measurements covered the entire global ocean (Parker et al. 1995). In particular, the Southern Ocean has generally not been monitored (Rayner et al. 2003). Therefore, the advent of reliable satellite SST retrieval systems has revolutionised the observation of SSTs in data sparse regions. Since the 1980s, satellite-borne infrared radiometers have been used to retrieve SST. The two most widely used instruments are the advanced very high resolution radiometer (AVHRR) and the European along-track scanning radiometers (ATSR) onboard the ERS1, ERS2 and ENVISAT earth observation satellites. These instruments measure infrared radiances from the Earth's surface in discrete wavelength channels (typically at 3.7 , 11 and $12\text{ }\mu\text{m}$). MSU are also used to measure surface temperatures, particularly for regions under persistent cloud cover such as the tropical Pacific (see Figure 2.3 and Section 2.4). However, their limitation is that the vertical and horizontal resolution is ~ 25 times coarser than AVHRR or ATSR.

The radiance relates to the temperature of the top few microns of the ocean (the skin), and this can differ from the bulk temperature sampled by traditional *in situ* measurements. Two physical processes contribute to the discrepancy between skin and bulk SSTs. The first is the skin effect, which arises as a result of conductive heat loss from the ocean, and generally results in a skin temperature up to 0.4°C cooler than the bulk temperature; the exact value depends on the net ocean-atmosphere heat flux and on wind speed. The second is the development of the diurnal thermocline, a thermally-stratified layer which builds up under conditions of strong insolation and low wind speed. This can result in a skin temperature up to 3°C warmer than the bulk temperature. These biases in satellite-derived SST data must be corrected to ensure consistency with the historical SST record. The

correction process is discussed in detail by Rayner et al. (2003).

Errors

Trenberth et al. (1992) listed five sources of error in monthly mean SSTs estimated from *in situ* measurements. All of these problems are present in the COADS data set. However, Bottomley et al. (1990) and Parker et al. (1995) discussed methods to minimise effectively the influence of these biases in the MOHSST6 data set:

- (i) Individual observations are not particularly accurate. The standard deviation between engine intake and bucket temperatures taken from the same ship was found to be 0.9°C (Saur 1963). Moreover, ship to ship observations made within 100 km and 6 hours had a standard deviation of differences $\sim 1.5^{\circ}\text{C}$ (Bernstein and Chelton 1985). Based on these studies, the expected standard error of an individual observation is 1°C .
- (ii) Incomplete sampling of the diurnal cycle can be significant. Cornillon and Strama (1985) show that at times when winds are light, diurnal heating can exceed 2°C . Typical values are much less than this and are generally of the order 0.2 to 0.3°C (e.g., Morrissey (1990) for north of Australia).
- (iii) Incomplete sampling of intramonthly variance other than the diurnal cycle and seasonal cycles (Fu et al. 1988).
- (iv) Incomplete sampling of intramonthly mean variance due to the seasonal cycle, which has typical amplitudes for SST of 3°C in the Southern Hemisphere and 5°C in the NH but only 1°C in the tropics.
- (v) SST gradients across a single grid box can cause errors in regions of strong gradients (e.g., Kuroshio and Gulf Stream currents). Minor changes in ship track across such regions can lead to substantial apparent anomalies in local SST. For large area averages the effects should be less since the random errors will tend to cancel each other out.

Satellite data are ‘tuned’ by regression against quality controlled drifting buoy data. The corrections are often quite large ($\sim 1^{\circ}\text{C}$) (Reynolds 1988), which brings into question the reliability of satellite estimates in regions where this calibration is not performed directly. While retrieval algorithms adjust for attenuation by tropospheric water vapour, they can be fooled, and the temperature of cloud tops can be mixed with SSTs. Aerosols emitted by volcanic eruptions can also lead to serious biases (Reynolds and Smith 1995).

Data set	Time period	Resolution	Coverage	Reference
ERSST2 ^a	1854 to 1997	2° × 2°	Global	Smith and Reynolds (2004a)
GISST ^b	1871 to 2003	1° × 1°	Global	Rayner et al. (1996)
HadISST1 ^c	1871 to present	1° × 1°	Global	Rayner et al. (2003)
Kaplan	1856 to 1991	5° × 5°	Near global	Kaplan et al. (1998)
NCEP OI.v2 ^d	1950 to present	1° × 1°	Global	Reynolds et al. (2002)

^a Extended Reconstructed Sea Surface Temperature (version 2).

^b Global Ice and Sea Surface Temperature.

^c UKMO Hadley Centre Ice and Sea Surface Temperature.

^d NCEP Optimal Interpolation analysis (version 2).

Table 2.3: Global analyses of sea surface temperature data.

Global gridded data sets

Table 2.3 summarises the five main global reconstructed SST data sets available in terms of their temporal coverage, spatial extent and resolution. While the details of the analyses vary, their reconstruction divides into three steps:

- (i) Bias correction of satellite data using high quality ship and buoy data.
- (ii) Estimation of spatial covariance structure of bias corrected SSTs.
- (iii) Interpolation of historical *in situ* data to match modern covariance structure.

Each SST reconstruction is a blend of *in situ* and satellite measurements and is therefore subject to the errors described above. Additional uncertainty arises from the particular interpolation procedure applied. Since no SST data set can be described as definitive, it is difficult to quantify these errors. Differences between analyses can be revealed by comparison to the reference data where available. However, such comparisons are not independent since much of the reference data will have been incorporated into the analyses. Therefore, differences are likely to be sensitive to the reconstruction method employed. The development of SST data sets remains an active area of research (e.g., Rayner et al. 2003, Smith and Reynolds 2004a).

Figure 2.4 shows the comparison of annual and near-decadal averages from five SST data sets over six different regions. Grid boxes partially covered by sea ice

are excluded from the area averaging. The time series shown are anomalies with respect to the 1961-90 mean. Good agreement between data sets is seen over the Gulf Stream region (panel a), the Baltic Sea (panel d) and the Pacific Niño 3.4 region (panel f). The Kaplan et al. (1998) data exhibit the greatest differences to the other data sets before 1940 over the Kuroshio region (panel b). The highest spread of SST values is observed over the Southern Ocean (panel e), which is the least well observed region.

Choice of SST data

Reconstructed SST data are chosen because of their continuous coverage and the rigorous quality control applied in their formulation. Of the reconstructed data, HadISST1 is favoured because of its 1° spatial resolution, temporal coverage back to 1871, lack of bias and weak trend with respect to the reference data (Rayner et al. 2003). However, because of the uncertainties outlined above, care must be taken when interpreting results originating from data sparse regions.

Sea ice data are incorporated into HadISST1 and are important for forcing atmospheric GCM simulations. However, the sea ice data are heterogeneous because sea ice has been observed using a variety of methods and in very different levels of detail throughout the historical record (Rayner et al. 2003). Furthermore, sea ice has a longer intrinsic time scale than SST and grid boxes with fractional ice cover can have detrimental effects on SST analyses. The present study employs SST data extensively to investigate seasonal-to-interannual predictability. Sea ice is therefore removed prior to computing all analyses to minimise these effects.

2.6 Snow cover

Background

Station measurements of snow cover have been made for over a century in parts of the Former Soviet Union, China and the United States (Brown 2000). However, the spatial coverage provided by these stations is generally limited to low eleva-

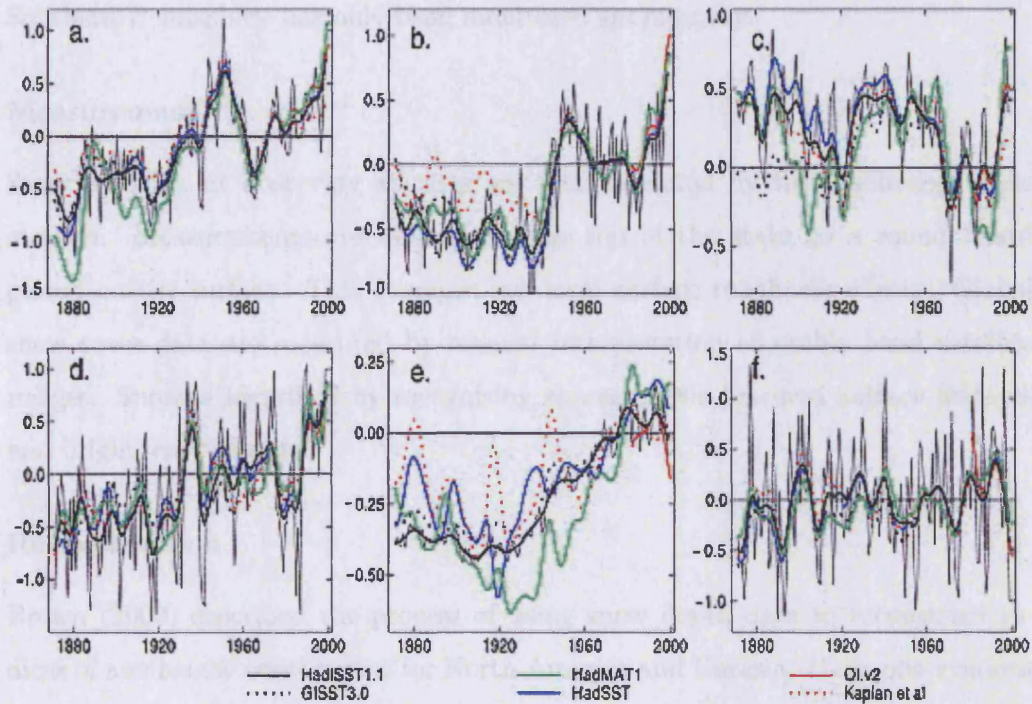


Figure 2.4: Smoothed annual SST anomalies averaged over (a) Gulf Stream region (35°N - 45°N , 70°W - 50°W); (b) Kuroshio region (30°N - 40°N , 125°E - 160°E); (c) Greenland region (50°N - 70°N , 70°W - 30°W); (d) Baltic Sea; (e) Southern Ocean south of 50°S ; (f) Pacific Niño 3.4 region (5°S - 5°N , 170°W - 120°W). Anomalies are calculated with respect to the 1961-90 mean and only ice-free grid boxes with data are used. A 21-point binomial (near-decadal) filter was applied to annual anomalies. Thin black curves are unsmoothed HadISST1 annual anomalies. “HadMAT1” refers to the Hadley Centre night time marine air temperature data set, “OI.v2” refers to Reynolds et al. (2002), “GISST3.0” refers to an updated version of Rayner et al. (1996), “HadSST” refers to Jones et al. (2001) and “Kaplan et al.” refers to Kaplan et al. (1998). After Rayner et al. (2003). Figure is © Crown Copyright 2003.

tion regions and to snow courses in mountainous regions (Robinson et al. 1993). An urban bias also exists which may give rise to an underestimation of coverage. Global snow cover data have only become available since the advent of remote sensing satellites. NH lands have been monitored continuously since the 1960s. The Southern Hemisphere has only been monitored sporadically.

Measurement

Snow amounts at observing stations are still measured by the traditional stake method. Measurements are taken from the top of the stake to a round board placed on the surface. This averages out local surface roughness effects. Global snow cover data are measured by manual interpretation of visible band satellite images. Snow is identified by recognising characteristic textured surface features and brightness patterns.

Reconstruction

Brown (2000) described the process of using snow depth data to reconstruct indices of areal snow cover extent for North America and Eurasia. Daily observations of snow depth from more than 500 stations with complete records were collected. These data were then interpolated to a $\sim 200\text{km}^2$ polar stereographic grid. The effect of topography was ignored because the station network remained fixed throughout the reconstruction period and the individual stations were at low elevation. Monthly snow cover values were derived by averaging the total snow covered (above a threshold depth of 2 cm) land area in the daily grids.

Errors

Robinson et al. (1993) listed four potential sources of error in satellite derived estimates of snow cover:

- (i) The inability to detect snow cover when solar illumination is low or when skies are cloudy.
- (ii) The underestimation of cover where dense forests mask the underlying snow.

- (iii) Ambiguities in the recognition and demarcation of patchy snow cover.
- (iv) Difficulties in discriminating snow from clouds in mountainous regions and in uniform lightly vegetated areas that have high surface brightness when snow covered.

Brown (2000) listed the main sources of error in the snow cover reconstruction:

- (i) Serious data gaps over China, Mongolia and Siberia meant that only data for March, April and October could be reconstructed for Eurasia.
- (ii) The number of stations with data over Eurasia was only one-third of that over North America and their locations were biased towards western Eurasia. However, this problem may not be serious because most of the interannual variability in Eurasian snow cover occurs over western Eurasia (Clark et al. 1999).
- (iii) Relaxation of the complete record criterion increased spatial coverage but introduced substantial errors into the snow cover indices ($\sim 10\%$ error by allowing just one missing year in the station records).
- (iv) There was some sensitivity to the chosen value of snow depth threshold. A value of 2 cm was used but other authors have used thresholds of 3 cm (Frei and Robinson 1998) and 2.5 cm (Frei et al. 2003).

Choice of snow cover data

The longest satellite based records available are provided by Rutgers University. These begin in 1966 but are only considered to be reliable since 1972 (Robinson et al. 1993). The charts are provided weekly on an 89×89 cell NH grid having a polar stereographic projection. Cell resolution ranges from $16,000\text{km}^2$ to $42,000\text{km}^2$. Only cells whose area is interpreted to be at least 50% snow covered are considered snow covered.

A secondary source of snow cover data are the indices of Brown (2000). The data are available (and considered reliable) for the period 1915-97 for North America and the period 1922-97 for Eurasia. Prior to 1971, the snow cover indices are derived entirely from station snow depth data, but after 1971 they are blended with the satellite observations. However, due to data sparsity, the reconstruction was only performed November-April for North America and in three months (March, April and October) for Eurasia.

Parameter	Data set	Temporal coverage	Spatial resolution	Source
MSLP	GMSLP2.1f	1871 to 1994	$5^{\circ} \times 5^{\circ}$	UKMO
MSLP	NCEP/NCAR	1948 to present	$2.5^{\circ} \times 2.5^{\circ}$	CDC ^a
NAO	CRU Index	1821 to present	-	CRU ^b
NAO	Hurrell	1865 to present	-	NCAR ^c
NAO	Hurrell MSLP PC1	1899 to present	-	NCAR
Temperature	CRUTEM2	1856 to present	$5^{\circ} \times 5^{\circ}$	CRU
Temperature	NCEP/NCAR	1948 to present	$2.5^{\circ} \times 2.5^{\circ}$	CDC
SST	HadISST1	1871 to present	$1^{\circ} \times 1^{\circ}$	UKMO
SST	NCEP/NCAR	1948 to present	$2.5^{\circ} \times 2.5^{\circ}$	CDC
Snow cover	NOAA Satellite	1972 to present	$2^{\circ} \times 2^{\circ}$	Rutgers GSL ^d
Snow cover	Brown	1922 to 1997	-	NSIDC ^e

^aUS Climatic Data Center.

^bUniversity of East Anglia, Climatic Research Unit.

^cUS National Center for Atmospheric Research.

^dRutgers University Global Snow Lab.

^eNational Snow and Ice Data Center.

Table 2.4: Summary of the main data sets used in this study.

2.7 Summary

Table 2.4 shows the main data sets used in this study. All other data not mentioned specifically in this table are taken from the NCEP/NCAR reanalysis (Kalnay et al. 1996). This includes wind and temperature fields on pressure levels and surface radiative fluxes. As described in Section 2.4, reanalysis data assimilate observations, satellite and coupled GCM data to produce complete global coverage. Therefore, caution should be exercised when interpreting results derived using reanalysis data over regions with few observing locations. Furthermore, few accurate climatological observations of surface radiative fluxes were made over land prior to satellite data becoming available. This means that these data in particular may suffer from inhomogeneities in the early record that cannot be verified against other observations.

Chapter 3

What is the Best Lagged Predictor of the Winter NAO?

3.1 Introduction

Bjerknes (1964) hypothesised that fluctuations in the winter NAO (henceforth NAO_{DJF}) index on interannual and longer time scales are associated with ocean-atmosphere interaction. Since then, observational and GCM investigations have confirmed the influence of the NAO_{DJF} circulation on the Atlantic Ocean. However, the question of whether — and to what extent — the ocean or land surface imparts a reciprocal influence on the atmosphere remains a topic of much debate (Bretherton and Battisti 2000, Drévilion et al. 2001, Marshall and Coauthors 2001). The presence of a lagged influence on the NAO_{DJF} would offer the prospect of seasonal to decadal winter climate prediction over the North Atlantic and adjacent continents (Paeth et al. 2003).

The aim of this Chapter is to examine “What is the best lagged predictor of the winter NAO?”. This is investigated by providing a standardised platform on which to evaluate and compare the predictive NAO_{DJF} skill arising from four previously published lagged NAO_{DJF} predictors and a new predictor. The Chapter is structured as follows. Section 3.2 outlines the chosen assessment periods, predictand NAO_{DJF} indices, predictors and the different observational data sets used to compute them. Here we also describe our hindcast procedure and the method for evaluating hindcast skill and its statistical significance. Results from the hindcast comparison are presented in Section 3.3. The stationarity of the predictive relationships is also discussed. Section 3.4 contains interpretation of the key findings

from these results and a detailed examination of the physical mechanisms linking the most skilful lagged predictor to the NAO_{DJF} . Section 3.5 presents a discussion of our findings and their implications for the broader topics of seasonal forecasting and twentieth century North Atlantic sector climate variability. Finally, a summary and conclusions are presented in Section 3.6.

3.2 Methodology

3.2.1 Data

NAO_{DJF} indices

Figure 3.1 shows the three NAO_{DJF} indices employed as predictands in this study. The selected indices are the CRU, Hurrell and MSLP PC1 NAO_{DJF} indices defined in Section 2.3. Despite being calculated using different methods and with different data sets, the three indices cross-correlate strongly over the period 1900–2001. The following are the Pearson product-moment correlation coefficients between the NAO_{DJF} indices: CRU/Hurrell $r = 0.88$, Hurrell/MSLP PC1 $r = 0.85$ and CRU/MSLP PC1 $r = 0.86$. However, as shown in Figure 3.1, there are quantitative and qualitative differences between the indices. The CRU index has a clear positive bias (100-year mean is 0.56) and has the highest 100-year variance (1.52 compared to 1.32 and 1.02 for the Hurrell and MSLP PC1 indices, respectively).

Previously published NAO_{DJF} predictors

The four previously published NAO_{DJF} predictors examined in this study are shown in Table 3.1 and are all taken from results published since 2001. They were selected for comparison because they are believed to encompass the highest currently observed seasonal predictive skill for the winter NAO.

The first NAO_{DJF} predictor (henceforth May SST (SVD)) is the time series of May North Atlantic (10°N–80°N) SST derived from a singular value decomposition (SVD) analysis with DJF mean 500 hPa geopotential height over the North Atlantic sector (Rodwell and Folland 2002). In the original analysis, the SVD procedure

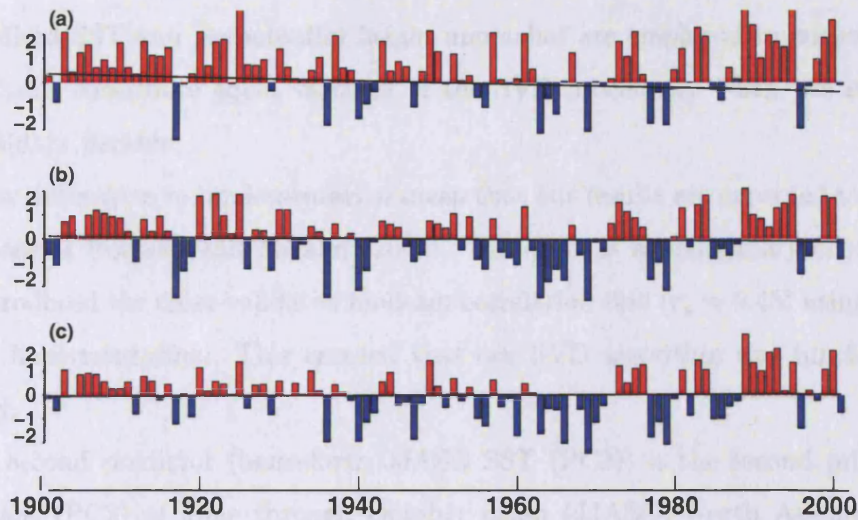


Figure 3.1: Temporal evolution of the (a) CRU, (b) Hurrell and (c) MSLP PC1 NAO_{DJF} indices 1900-2001 in standardised units. Solid line indicates linear trend as determined by a least squares fit.

was performed in ‘cross-validated’ mode, where each year in turn was removed from the data set and the SVD analysis repeated. Time series were calculated by projecting onto the data for each removed year the loading patterns calculated using the remaining data. The resulting SST time series was called the ‘predicted’ NAO series and the corresponding geopotential height time series the ‘observed’ NAO series. The correlation skill was taken as the correlation between these two time series.

Our implementation of the SVD analysis differs from that employed by Rodwell and Folland (2002). First, in accordance with other analyses in this study, we employ a five-year block elimination to the SVD procedure. This means that we remove each year in turn and additionally two years either side. This mitigates against serial correlation artificially inflating the covariance between the SST and geopotential height time series. Second, we employ the SST time series from the SVD as the ‘predicted’ time series and the CRU, Hurrell and MSLP PC1 NAO_{DJF} indices as the ‘observed’ time series. Third, we remove sea ice from the SST fields prior to performing the SVD (sea ice was retained in the original analysis). Fourth,

standardised SST and geopotential height anomalies are employed to ensure that all grid cells contribute equal variance in the SVD procedure. Fifth, we employ different data periods.

These differences in implementation mean that our results are expected to differ from those of Rodwell and Folland (2002). However, as a preliminary check, we first reproduced the cross-validated hindcast correlation skill ($r_s = 0.45$) using their original implementation. This ensured that our SVD algorithm was functioning correctly.

The second predictor (henceforth JJASO SST (PC2)) is the second principal component (PC2) of June through October mean (JJASO) North Atlantic (0°-65°N) SST (Saunders and Qian 2002). This is slightly the stronger of two lagged JJASO SST predictive modes employed by Saunders and Qian (2002). The loading pattern of the second EOF is derived using standardised anomalies of SST to ensure that all grid cells contribute equal variance in the principal component analysis. This pattern (see Saunders and Qian (2002) their Figure 1a) features a ring of SST anomalies around an opposite-signed centre off Newfoundland. Radcliffe and Murray (1970) were the first to identify a relationship between autumn SST anomalies off Newfoundland and the atmospheric circulation during subsequent months over western Europe.

The third and fourth NAO_{DJF} predictors employ the monthly mean areal extent of snow cover, with observations considered reliable since 1972. The third predictor is snow cover over Eurasia in October (Saito et al. 2001). The fourth predictor is the June-July (JJ) mean snow cover for the entire Northern Hemisphere (Saunders et al. 2003).

New NAO_{DJF} predictor

Alongside the four previously published predictors outlined above, we also examine a new NAO_{DJF} predictor that is an index of the zonal gradient in June-July (JJ) Northern Hemisphere (NH) subpolar 2 m air temperature. This index and its link

Lagged Predictor	Assessment Period	Domain Area	NAO _{DJF} Index	Reference
May SST (SVD)	1948-1998	90W-40E, 10N-80N	Z500 SVD	Rodwell and Folland (2002)
JJASO SST (PC2)	1950-2001	100W-0E, 0N-65N	CRU/CPC/PC1	Saunders and Qian (2002)
Oct EU Snow Cover	1972-2000	n/a	AO*	Saito <i>et. al</i> (2001)
JJ NH Snow Cover	1972-2002	n/a	CRU/CPC**/PC1	Saunders <i>et. al</i> (2003)

* Arctic Oscillation index (Thompson and Wallace 1998). ** Climate Prediction Center, U.S.

Table 3.1: Summary of current NAO seasonal prediction studies examined in this Chapter.

to NAO_{DJF} was introduced by Saunders et al. (2003). The index is defined as

$$\Delta_x T_{60N-70N} = \frac{NA + EU}{2} - SG, \quad (3.1)$$

where NA, EU and SG refer to subpolar (60°N-70°N) 2 m air temperatures over centres in North America, Eurasia and southern Greenland, respectively. These centres are shown in Figure 3.2 and correspond to where gridded JJ air temperature is correlated most significantly with contemporaneous NH snow cover. The two negative centres are a direct thermal response to anomalous snow cover, while the positive centre is a contemporaneous teleconnection.

Gridded 2 m air temperature data are available back to 1900 from the CRUTEM2 data set (Jones and Moberg 2003) for NA, EU and SG. We therefore calculate the $\Delta_x T_{60N-70N}$ index (henceforth ΔT) for the period 1900-2001 and employ it as an NAO_{DJF} predictor. The summer period is selected because during the summer months the relationship peaks between the ΔT index and NAO_{DJF} (Saunders et al. 2003). Figure 3.3 shows that the link is strongest in JJ using both the CRUTEM2 and NCEP/NCAR reanalysis data (Kalnay et al. 1996) for the 50- and 30-year assessment periods. However, for the 100-year period the link is significant over an extended summer period May through September (MJJAS). Therefore, we also evaluate the NAO_{DJF} predictive skill of the longer MJJAS mean ΔT .

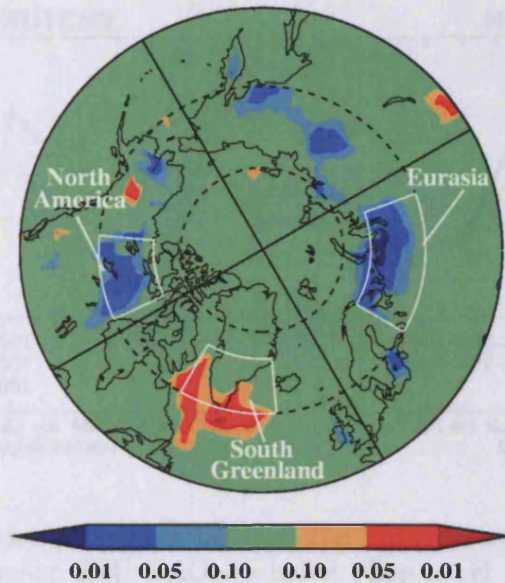


Figure 3.2: The correlation pattern significance between detrended time series of June-July Northern Hemisphere snow cover extent and gridded June-July 2 m air temperature. Significances are corrected for serial correlation with lags out to 15 years included. Colour scale denotes where correlation is positive (reds) or negative (blues). After Saunders et al. (2003).

Predictor data sets

As outlined above, this work reexamines the results of previously published studies to standardize and compare NAO_{DJF} predictive skill. Wherever possible, we employ the same data sets used in the original works. However, as we also extend the analysis period back in time to 1900, it was necessary to use alternative data sets to provide data for this longer period. The data sets used in this study are listed in Table 3.2

3.2.2 Trends

Throughout this thesis, the influence of linear (i.e., monotonic) trends on correlation and regression analyses is considered. When correlation or linear regression are used to investigate predictability, the analysis is computed using both raw data (i.e., with trends included) and detrended data (with a linear trend removed). This quantifies the influence of monotonic trends on predictability. A climate process Y can be

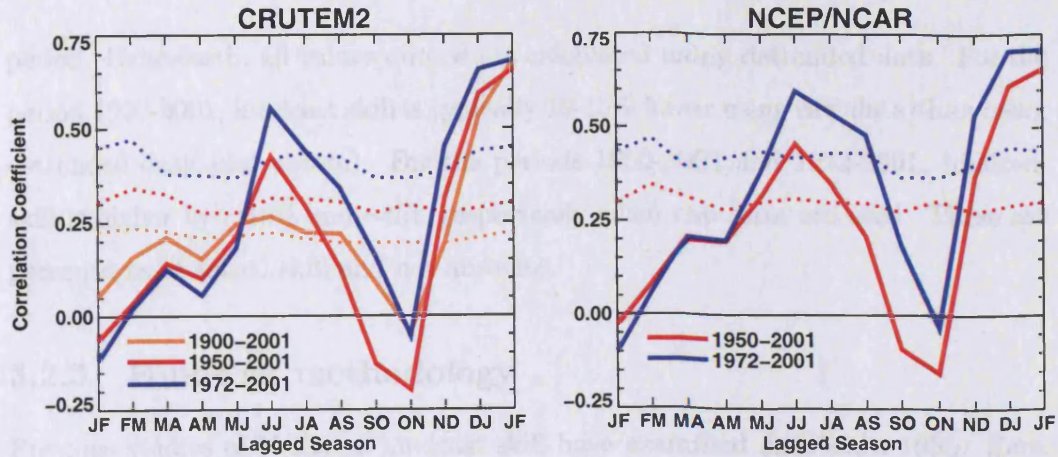


Figure 3.3: Correlation coefficient between the lagged seasonal ΔT temperature index and the upcoming CRU NAO_{DJF} index. Top panel shows correlations for the CRUTEM2 ΔT index and bottom panel is for NCEP/NCAR data. Dark lines denote correlation coefficient and faint lines denote 5% significance level corrected for serial correlation.

decomposed into

$$Y(t) = T(t) + \varepsilon(t), \quad (3.2)$$

where T is its deterministic (predictable) component, ε is the unpredictable (noise) component and t is time. Assuming T is a linear trend, it can be modelled using the relation

$$T(t) = \beta_0 + \beta_1 t, \quad (3.3)$$

where the parameters β_0 and β_1 are the intercept and slope of the trend line, respectively. The latter represents the trend magnitude. These coefficients are estimated using an ordinary least squares regression fit of Y against t . The statistical significance ($p < 0.05$) of the trend is estimated using the Mann-Kendall test (e.g., Hisdal et al. 2001). This test is conducted against H_0 that the time series contains no trend.

In this Chapter, a linear trend is removed from the time series of all predictors and NAO_{DJF} indices prior to performing the hindcast procedure. This ensures that trends do not influence either the model formulation or the skill assessment. The trend is calculated as a least-squares regression fit over each full assessment

period. Henceforth, all values quoted are calculated using detrended data. For the period 1900-2001, hindcast skill is typically 10-15% lower using raw data than using detrended data (not shown). For the periods 1950-2001 and 1972-2001, hindcast skill is higher by $\sim 10\%$ and $\sim 5\%$, respectively when raw data are used. These are percentages of actual skill and not absolute.

3.2.3 Hindcast methodology

Previous studies of NAO_{DJF} hindcast skill have examined data since 1950. Here, we assess hindcast skill, where possible, over the entire twentieth century. Specifically, NAO_{DJF} predictability is assessed for the three periods 1900-2001, 1950-2001 and 1972-2001, with the 1972-2001 period corresponding to the interval of reliable satellite snow cover data (Robinson et al. 1993). These three assessment periods also allow the stationarity in NAO_{DJF} hindcast skill to be examined.

The predictive skill of the selected NAO_{DJF} predictors is computed for each assessment period using cross-validated (Michaelsen 1987) hindcasts with block elimination (WMO 2002) and (except for the May SST (SVD) predictor) linear regression models. Linear least squares regression is valid if the predictor, predictand (NAO_{DJF}) and hindcast error time series are all normally (Gaussian) distributed (Wilks 1995). The Lilliefors goodness-of-fit test (Wilks 1995) is used to examine normality. The null hypothesis (H_0) is that the data are drawn from a population with a theoretical Gaussian probability density function (PDF). For all predictor variables and the three NAO_{DJF} indices we could not reject H_0 using the Lilliefors test and therefore conclude that all distributions are not significantly different from Gaussian.

The hindcast model to predict the NAO_{DJF} index \hat{y} in a given year t using predictor x takes the form

$$\hat{y}_t = \beta_0 + \beta_1 x_t + e_t \quad (3.4)$$

where the coefficients β_0 and β_1 are determined by an ordinary least squares regression and e_t is the residual. A five-year block elimination is employed, which means that for a given predictand year t , data for the years $t-2$, $t-1$, t , $t+1$ and

Chapter 3 What is the Best Lagged Predictor of the Winter NAO?

Predictor	1900-2001	1950-2001	1972-2001
May SST (SVD)	GMSLP2.1f*, HadISST1	NNR Z500 & SST, HadISST1	NNR Z500 & SST, HadISST1
JJASO SST (PC2)	HadISST1	NNR SST, HadISST1	NNR SST, HadISST1
Oct EU Snow Cover	Brown/Rutgers**	Brown/Rutgers	Brown/Rutgers
JJ NH Snow Cover	-	-	Rutgers

*Data available 1900-1994. **Data available 1922-2001

Table 3.2: Data sets used to compute NAO_{DJF} predictors in this study. NNR denotes NCEP/NCAR reanalysis data.

$t+2$ are removed from the regression. This mitigates against inflation of hindcast skill due to serial correlation in the predictor and/or predictand time series. Furthermore, predictors derived from principal components (PCs) are also subjected to block elimination in their formulation. This ensures that the PCs do not contain information from years that are eliminated from the regression procedure. For each predictor a time series of hindcasted NAO_{DJF} values is calculated whose skill is verified against the corresponding observed NAO_{DJF} series.

Skill assessment

In validating the hindcast skill from each predictor, we employ two skill measures. These are the correlation skill (r_s) and the mean-squared skill score ($MSSS$) against a simple prediction of climatology. The former is the Pearson product-moment correlation coefficient between the observed and predicted NAO_{DJF} time series. The latter is the measure recommended by the World Meteorological Organization for verification of deterministic seasonal forecasts (WMO 2002) and is defined as

$$MSSS_j = (1 - \frac{MSE_j}{MSE_{cj}}) \times 100\% \quad (3.5)$$

where MSE_j is the mean-squared error of the hindcasts and MSE_{cj} is the mean-squared error for climatology. The climatology used in this study is the long-term mean for each assessment period.

The statistical significance of each skill value is estimated using a randomised resampling method (Manly 1997). The observed and hindcast NAO_{DJF} time series are both randomly shuffled (with replacement) to create two new synthetic time

series drawn from the same populations as the original time series. The length of the two synthetic time series is reduced to equal the ‘effective’ number of degrees of freedom (discussed below) between the original time series. Skill values are calculated for the two randomised time series and this shuffling/resampling process is repeated in Monte Carlo fashion 25,000 times. The statistical significance is determined by the number of times that the skill value from the random data exceeds the original observed skill value. Our null hypothesis (H_0) is that there is no link between a given predictor and NAO_{DJF} and that any observed skill is achieved by random chance. The *a priori* threshold of statistical significance is set at 0.05, which represents the probability that H_0 is incorrectly rejected (Type I error). Our test is one-tailed as only positive values of correlation skill and $MSSS$ will lead to a rejection of H_0 (Wilks 1995).

The influence of serial correlation between two time series reduces the number of statistical degrees of freedom (DOF) available to the analyst. Failure to correct for this reduction increases the chance of Type I errors. Throughout this thesis we correct for serial correlation using the method of Davis (1976), which involves calculating the *effective* DOF between the time series that are being correlated. The effective DOF is $N' = N/\tau$, where N is the actual DOF and τ is the reduction factor

$$\tau = \sum_{i=-\infty}^{\infty} \rho_X(i)\rho_Y(i) \approx 1 + 2 \sum_{i=1}^n \rho_X(i)\rho_Y(i) \quad (3.6)$$

and ρ is the autocorrelation function over n lagged years (Eqn.4.4). The value of n is typically one-half the length of the time series. Analysis of all predictors shows that the May SST (SVD) predictor has the highest serial correlation with a lag-1 correlation coefficient of 0.53 and a mean over ten lagged years of 0.38. The other predictors have lag-1 coefficients between 0.20 and 0.30 and a mean over ten years of around 0.10.

Field significance

Statistical analyses of gridded climate fields suffer from problems associated with temporal autocorrelation (discussed above) and also problems associated with spa-

tial autocorrelation. The stronger the spatial autocorrelation in a field, the lower the independence between sampling locations, which reduces the spatial DOF. This increases the chances of finding coherent regions of, for example, significant correlation because large numbers of non-independent statistical tests are performed on a finite set of data locations. The coherent regions imply a strong physical association over a large area between the variables being correlated. However, because of the lack of independence between the locations, the region may in fact be a series of interrelated Type I errors.

Field significance and the number of spatial DOF in a field are estimated using the two-stage method of Livezey and Chen (1983). First, the number of locally significant tests that are required for field significance ($p < 0.05$) is computed using the binomial distribution, assuming all locations are independent. If the observed number of locally significant tests does not exceed the number required assuming independence, then the field is not significant. Second, a Monte Carlo resampling method is used to repeat the original test but using Gaussian random data, i.e., violating H_0 deliberately each time. This provides a PDF of the expected number of locally significant tests occurring in the field by random chance. The upper bound of spatial DOF in the observed field is computed by comparing the $p = 0.05$ level on the PDF with the original binomial distribution.

3.3 NAO hindcast skill

Table 3.3 shows the NAO_{DJF} cross-validated hindcast skill (and significance) for each predictor against the three NAO_{DJF} indices over all three assessment time periods. The analysis of these results based on assessment time period now follows.

3.3.1 1900-2001

For the 100-year period, significant hindcast skill is found for all three NAO_{DJF} indices using both ΔT index predictors. The long summer (MJJAS) mean produces the most significant skill (a 6-9% improvement over a prediction of climatology) and

Period	Lagged Predictor	Data set	CRU NAO _{DJF}				Hurrell NAO _{DJF}				MSLP PC1 _{DJF}			
			<i>r</i>	(<i>p</i>)	<i>MSSS</i>	(<i>p</i>)	<i>r</i>	(<i>p</i>)	<i>MSSS</i>	(<i>p</i>)	<i>r</i>	(<i>p</i>)	<i>MSSS</i>	(<i>p</i>)
1900-2001	May SST (SVD)*	Had	0.11	(0.19)	0	(-)	0.15	(0.10)	2	(0.07)	0.15	(0.14)	1	(0.10)
	JJASO SST (PC2)	Had	0.13	(0.11)	1	(0.08)	0.15	(0.09)	1	(0.07)	0.09	(0.06)	0	(-)
	JJ $\Delta_x T_{60N-70N}$	CRUT	0.18	(0.05)	2	(0.04)	0.24	(0.02)	5	(0.02)	0.21	(0.04)	4	(0.03)
	MJJAS $\Delta_x T_{60N-70N}$	CRUT	0.26	(0.01)	6	(0.01)	0.31	(<0.01)	9	(<0.01)	0.28	(0.01)	7	(<0.01)
	Oct EU Snow Cover [#]	Brn/Rut	0.13	(0.12)	1	(0.09)	0.08	(0.25)	0	(-)	0.28	(0.01)	7	(<0.01)
1950-2001	May SST (SVD)	Had	0.07	(0.35)	0	(-)	0.16	(0.17)	1	(0.11)	0.12	(0.25)	0	(-)
	May SST (SVD)	NCEP	0	(-)	0	(-)	0.11	(0.28)	0	(-)	0.03	(0.44)	0	(-)
	JJASO SST (PC2)	Had	0.24	(0.05)	5	(0.03)	0.25	(0.05)	6	(0.03)	0.17	(0.12)	1	(0.09)
	JJASO SST (PC2)	NCEP	0.24	(0.05)	5	(0.03)	0.24	(0.05)	5	(0.03)	0.18	(0.10)	2	(0.07)
	JJ $\Delta_x T_{60N-70N}$	CRUT	0.33	(0.02)	10	(0.01)	0.27	(0.05)	6	(0.03)	0.33	(0.03)	10	(0.02)
	JJ $\Delta_x T_{60N-70N}$	NCEP	0.46	(<0.01)	21	(<0.01)	0.36	(0.02)	12	(0.01)	0.46	(<0.01)	21	(0.01)
	JJ $\Delta_x T_{60N-70N}$	NCEP/CRUT	0.44	(<0.01)	19	(<0.01)	0.36	(0.02)	12	(0.01)	0.45	(<0.01)	20	(<0.01)
	MJJAS $\Delta_x T_{60N-70N}$	CRUT	0.21	(0.10)	4	(0.07)	0.27	(0.06)	7	(0.04)	0.33	(0.03)	11	(0.02)
	MJJAS $\Delta_x T_{60N-70N}$	NCEP	0.33	(0.02)	10	(0.01)	0.32	(0.04)	9	(0.02)	0.43	(0.01)	18	(0.01)
	MJJAS $\Delta_x T_{60N-70N}$	NCEP/CRUT	0.29	(0.04)	8	(0.03)	0.33	(0.03)	10	(0.02)	0.42	(0.01)	17	(0.01)
1972-2001	Oct EU Snow Cover	Brn/Rut	0	(-)	0	(-)	0	(-)	0	(-)	0.15	(0.15)	0	(-)
	JJ $\Delta_x T_{60N-70N}$	CRUT	0.54	(<0.01)	29	(<0.01)	0.57	(<0.01)	32	(<0.01)	0.51	(0.01)	25	(<0.01)
	JJ $\Delta_x T_{60N-70N}$	NCEP	0.58	(<0.01)	34	(<0.01)	0.56	(<0.01)	31	(<0.01)	0.54	(0.01)	29	(<0.01)
	JJ $\Delta_x T_{60N-70N}$	NCEP/CRUT	0.59	(<0.01)	35	(<0.01)	0.57	(<0.01)	33	(<0.01)	0.55	(0.01)	30	(<0.01)
	MJJAS $\Delta_x T_{60N-70N}$	CRUT	0.45	(0.03)	20	(0.02)	0.49	(0.02)	23	(0.01)	0.45	(0.03)	19	(0.02)
	MJJAS $\Delta_x T_{60N-70N}$	NCEP	0.51	(0.01)	26	(<0.01)	0.47	(0.01)	22	(0.01)	0.50	(0.02)	24	(0.01)
	MJJAS $\Delta_x T_{60N-70N}$	NCEP/CRUT	0.48	(0.02)	22	(0.01)	0.47	(0.02)	22	(0.01)	0.47	(0.02)	22	(0.01)
	Oct EU Snow Cover	Rut	0.26	(0.11)	6	(0.07)	0.07	(0.39)	0	(-)	0.26	(0.11)	5	(0.08)
	JJ NH Snow Cover	Rut	0.53	(0.01)	28	(<0.01)	0.51	(0.01)	24	(<0.01)	0.53	(0.01)	27	(<0.01)

*Data available 1900-1994. #Data available 1922-2001.

Table 3.3: Skill values from cross-validated NAO hindcasts. Data set abbreviations: Had = HadISST, CRUT = CRUTEM2, Brn/Rut = Combined Brown and Rutgers snow cover index, NCEP = NCEP/NCAR Reanalysis, Rut = Rutgers. *r* is the correlation skill value and *MSSS* is the percentage improvement in mean-square skill score over climatology. *p* is the probability that the observed skill value was obtained by random chance, with values shown in brackets. Bold type denotes skill values significant at less than 5% as determined by a Monte Carlo resampling test with 25,000 iterations.

the JJ mean results in a 2-5% improvement. In terms of correlation skill, the values range from 0.26 to 0.31 for MJJAS, which explains 7-10% of the variance in the NAO_{DJF} time series over 100-years. Significant skill is achieved using October Eurasian snow cover (data available 1922-2001) to predict the MSLP PC1 NAO_{DJF} index. However, the poor performance of this predictor against the two other NAO_{DJF} indices reduces confidence in this result. The May SST (SVD) (data available 1900-1994) and the JJASO SST (PC2) predictors both show positive skill over this period but the skill values are not statistically significant. Aside from October Eurasian snow cover there is little variability in predictor performance against NAO_{DJF} index. The two ΔT indices show highest skill in predicting the Hurrell NAO_{DJF} index but the differences are typically only around 10% of the total skill.

3.3.2 1950-2001

For the 50-year assessment period, statistically significant skill is achieved using the JJASO SST (PC2) predictor and both ΔT indices. The JJASO SST (PC2) predictor shows r_s values of 0.17-0.25 (3-6% variance explained) and $MSSS$ values of 1-6%. However, the skill predicting the MSLP PC1 NAO_{DJF} index is not statistically significant. The May SST (SVD) predictor exhibits no significant skill over this period against any of the NAO_{DJF} indices. In contrast to the 1900-2001 period, it is the JJ ΔT rather than the MJJAS ΔT that is most skilful for all three NAO_{DJF} indices. The former has maximum $r_s = 0.46$ and $MSSS = 21\%$, whereas for the MJJAS index the maximum skill values are $r_s = 0.43$ and $MSSS = 18\%$. The CRUTEM2 ΔT indices show skill 20-30% lower in this assessment period than the NCEP/NCAR ΔT indices. Calculating ΔT using the time-mean of the CRUTEM2 and NCEP/NCAR data sets produces skill levels close to those achieved using NCEP/NCAR alone. October Eurasian snow cover produces no skill against all three NAO_{DJF} indices. In terms of skill performance with NAO_{DJF} index, the JJASO SST (PC2) predictor performs best against the CRU and Hurrell indices. The JJ ΔT predictor shows highest skill against the CRU and MSLP PC1 NAO_{DJF}

indices, whereas for the MJJAS ΔT predictor the MSLP PC1 NAO_{DJF} index stands out against the other two.

3.3.3 1972-2001

Highest hindcast skill in this period is achieved using the JJ ΔT index and the areal extent of JJ NH snow cover. The JJ ΔT index shows the highest skill values (maximum $r_s = 0.59$ and $MSSS = 35\%$) this skill showing little sensitivity to NAO_{DJF} index or choice of temperature data set. JJ NH snow cover also shows highly significant skill (maximum $r_s = 0.53$ and $MSSS = 28\%$) against all three NAO_{DJF} indices. As in the 50-year assessment, the link between ΔT and the NAO_{DJF} indices is strongest for the JJ period rather than for the long summer (MJJAS) period. The combined CRUTEM2-NCEP/NCAR mean JJ ΔT index shows skill slightly higher than that achieved with each single data set JJ ΔT index. For MJJAS ΔT the skill lies between that achieved using CRUTEM2 and NCEP/NCAR data individually. Neither of the SST-based predictors or the October Eurasian areal snow cover extent produce significant skill for the 30-year assessment period. However, further analysis shows that significant hindcast skill is found during this period when the principal component for the JJASO SST (PC2) predictor is calculated using data extending over the period 1950-2001 (not shown) rather than 1972-2001 as here.

3.3.4 Stationarity

The increase in NAO_{DJF} predictive skill observed using the summer ΔT indices between the 1900-2001 assessment period and the 1972-2001 period suggests that the relationship between ΔT and the NAO_{DJF} is nonstationary. Therefore, the modest skill seen over the full 100-years could be drawn entirely from the 1972-2001 skilful period. To test this hypothesis we first perform a rolling cross-correlation analysis for both the JJ and MJJAS ΔT indices against the CRU NAO_{DJF} index for all possible 30-year periods 1900-2001 (Figure 3.4). A linear trend is removed from both time series prior to each correlation value being computed. The pattern of correlation variability is dominated for both indices by a low-frequency oscillation

with a period of around 60-years. The correlations have two distinct maxima in the 30-year periods beginning at about 1905 and 1965. The MJJAS correlation peaks for both maxima at around $r = 0.5$, whereas the two JJ peaks are at $r = 0.1$ and $r = 0.6$, respectively.

The analysis was repeated using high- and low-pass filtered NAO_{DJF} data to assess the influence of low-frequency variability on the rolling correlation (not shown). The high-pass filtered data show a similar correlation pattern to that in Figure 3.4, whereas the low-pass filtered data deviate from the original pattern during the period 1920-1950. This suggests that the multidecadal cycle in the original correlation pattern is more representative of the high-frequency than the low-frequency variations in the CRU NAO_{DJF} index.

Despite the coincidence between periods of high correlation and periods of strongly positive NAO_{DJF} index in the early and late twentieth century, the ΔT indices perform equally well in predicting above or below median NAO_{DJF} events. Hindcasts using MJJAS ΔT are correctly above and below median in 54% of the observed above and below median CRU NAO_{DJF} seasons 1900-2001. For the period 1950-2001 the JJ ΔT index is correct in 64% of cases, while for 1972-2001 the figure is 71%. We therefore conclude that the hindcast skill from the summer ΔT indices is nonstationary and this lack of stationarity is not linked to multidecadal variability in the NAO_{DJF} indices.

3.3.5 Data set dependence

Table 3.3 shows that, in general, the NAO_{DJF} hindcast skill exhibits little sensitivity to the choice of SST or 2 m air temperature data set. Specifically, the JJASO SST (PC2) predictor produces almost identical skill against the three NAO_{DJF} indices whether it is calculated using HadISST or NCEP/NCAR data 1950-2001. The largest skill sensitivity to data set is observed for the ΔT indices calculated using CRUTEM2 and NCEP/NCAR data 1950-2001. Using CRUTEM2 data, the JJ ΔT predictor explains 11% of the variance in the CRU and MSLP PC1 NAO_{DJF} indices compared with 21% using NCEP/NCAR data. The figures are 7% and

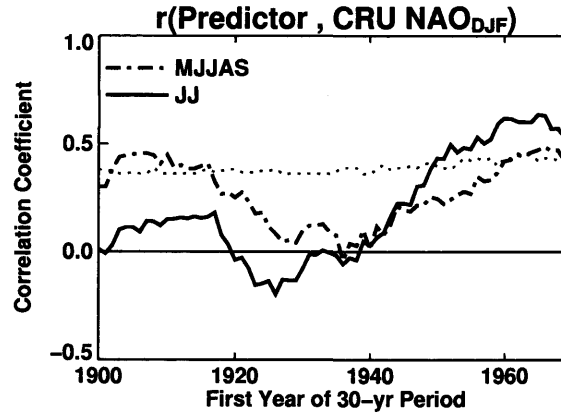


Figure 3.4: Correlation coefficients between the JJ and MJJAS ΔT indices and the upcoming CRU NAO_{DJF} index for running 30-year windows commencing 1900 to 1971. Faint dashed line indicates 5% significance level corrected for serial correlation.

13% for the Hurrell NAO_{DJF} index, respectively. A possible explanation for this sensitivity comes from the spatial coverage of the CRUTEM2 data set. Unlike the NCEP/NCAR data, the CRUTEM2 data are land-based and do not include data over the ocean. Since more than 50% of the southern Greenland ΔT region lies over the ocean and it has twice the weighting of the Eurasian and North American regions, the southern Greenland region is likely to contribute the largest errors in ΔT . Thus, the CRUTEM2 data may contain larger ΔT index errors than the NCEP/NCAR data.

3.4 Summer ΔT influence on upcoming NAO_{DJF}

3.4.1 Relationship between NAO_{DJF} lagged predictors

Quantifying the links between the different NAO_{DJF} lagged predictors will help to clarify the physical basis for the lagged predictability. Table 3.4 shows the cross-correlation matrix for all predictors over the three assessment periods. For the 100-year period, the SST predictors, JJ ΔT and MJJAS ΔT are all significantly correlated ($p < 0.05$). October Eurasian snow cover is only linked significantly to JJASO SST (PC2). The summer ΔT indices explain 27-29% of the 100-year

Period	Lagged Predictor	Data set	May		JJASO		JJ		MJJAS		Oct	JJ
			SST (SVD)		SST (PC2)		$\Delta_x T_{60N-70N}$		$\Delta_x T_{60N-70N}$		EU Snow	NH Snow
			Had	NCEP	Had	NCEP	CRUT	NCEP	CRUT	NCEP	Brn/Rut	Rut
1900-2001	May SST (SVD)	Had	-	-	0.30	-	0.30	-	0.35	-	0.07	-
	JJASO SST (PC2)	Had	-	-	-	-	0.54	-	0.52	-	0.23	-
	JJ $\Delta_x T_{60N-70N}$	CRUT	-	-	-	-	-	-	0.80	-	0.04	-
	MJJAS $\Delta_x T_{60N-70N}$	CRUT	-	-	-	-	-	-	-	-	0.02	-
	Oct EU Snow	Brn/Rut	-	-	-	-	-	-	-	-	-	-
1950-2001	May SST (SVD)	Had	-	0.78	0.35	0.56	0.26	0.47	0.29	0.41	0.04	-
	May SST (SVD)	NCEP	-	-	0.37	0.42	0.23	0.50	0.40	0.53	0.00	-
	JJASO SST (PC2)	Had	-	-	-	0.88	0.52	0.49	0.50	0.44	0.26	-
	JJASO SST (PC2)	NCEP	-	-	-	-	0.48	0.59	0.42	0.47	0.33	-
	JJ $\Delta_x T_{60N-70N}$	CRUT	-	-	-	-	-	0.69	0.82	0.58	0.01	-
	JJ $\Delta_x T_{60N-70N}$	NCEP	-	-	-	-	-	-	0.58	0.82	0.26	-
	MJJAS $\Delta_x T_{60N-70N}$	CRUT	-	-	-	-	-	-	-	0.77	0.07	-
	MJJAS $\Delta_x T_{60N-70N}$	NCEP	-	-	-	-	-	-	-	-	0.28	-
	Oct EU Snow	Brn/Rut	-	-	-	-	-	-	-	-	-	-
1972-2001	JJ $\Delta_x T_{60N-70N}$	CRUT	0.33	0.39	0.52	0.50	-	0.74	0.84	0.64	0.02	-0.51
	JJ $\Delta_x T_{60N-70N}$	NCEP	0.47	0.52	0.48	0.46	-	-	0.69	0.87	0.32	-0.67
	MJJAS $\Delta_x T_{60N-70N}$	CRUT	0.25	0.29	0.35	0.28	-	-	-	0.83	0.03	-0.55
	MJJAS $\Delta_x T_{60N-70N}$	NCEP	0.34	0.36	0.29	0.24	-	-	-	-	0.27	-0.58
	Oct EU Snow	Rut	0.16	0.15	0.22	0.29	-	-	-	-	-	0.20
	JJ NH Snow	Rut	0.29	0.26	0.43	0.40	-	-	-	-	-	-

Table 3.4: Matrix of cross correlations between predictors for all three assessment periods. Values are absolute Pearson product-moment correlation coefficients. Bold type denotes values significant at less than 5% as determined by a Monte Carlo resampling test with 25,000 iterations. Data set abbreviations are as in Table 3.3.

variance in the JJASO SST PC2 predictor. The links between the JJ and MJJAS ΔT indices and the JJASO SST (PC2) predictor also remain significant for the 1950-2001 period. October Eurasian snow cover is uncorrelated with any of the other predictors over this period. For the period 1972-2001, the correlation between the NCEP/NCAR JJ ΔT index and JJ NH snow cover is $r = -0.67$. This value is lower ($r = -0.51$) for the CRUTEM2 JJ ΔT index but remains significant. The JJ ΔT index (but not the MJJAS ΔT index) is related significantly to the JJASO SST (PC2) predictor. Also, JJ NH snow cover is not correlated significantly with the SST predictors. However, as before, if the JJASO SST (PC2) predictor is calculated using data over the period 1950-2001, the JJ NH snow cover predictor is correlated significantly to JJASO SST (PC2). October Eurasian snow cover is, once again, uncorrelated with any other predictor, including prior JJ NH snow cover.

3.4.2 Role of prior winter NAO_{DJF}

While a significant relationship between two NAO_{DJF} predictors could indicate that both predictors capture all or part of the same process related to the upcoming NAO_{DJF} , the relationship could also indicate an NAO_{DJF} -related influence from the previous winter on both predictors. For example, the May SST (SVD) and JJASO SST (PC2) patterns may arise from the persistence of the previous winter's NAO circulation over the North Atlantic. To test this possibility we reverse the hindcast procedure and use the lag-1 NAO_{DJF} to predict the May SST (SVD), JJASO SST (PC2) and JJ ΔT time series. Table 3.5 shows that there is a significant ($p < 0.05$) link between the prior winter NAO and upcoming May SST (SVD) pattern for the 100-year but not for the 50-year assessment period. There is also a link to the JJASO SST (PC2) predictor but this is only significant (for both the 100- and 50-year periods) when using the lag-1 Hurrell NAO_{DJF} index. The JJ ΔT index exhibits significant skill against the three prior NAO_{DJF} indices 1900-2001 with values slightly lower than for the May SST (SVD) predictor. Significant skill is not observed for the JJ ΔT index for the period 1950-2001. For the 1972-2001 period of highest NAO_{DJF} predictability (Table 3.3), neither the JJ ΔT index nor the two

SST indices are significantly predicted by any prior NAO_{DJF} index. In summary, Table 3.5 shows that there is little influence from the previous winter NAO on the predictors. This is consistent with each NAO_{DJF} index having an insignificant lag-1 autocorrelation over each assessment period (Table 3.5).

3.4.3 Role of lagged snow cover

Table 3.3 shows that October Eurasian snow cover has no predictive skill for NAO_{DJF} when assessed over all years. However, Cohen and Saito (2003) propose that autumn Eurasian snow cover influences Northern Hemisphere winter climate only during certain years. Such years are termed ‘Type-A’ because their atmospheric evolution resembles the AO pattern (Thompson and Wallace 1998). Conversely, winters that evolve with an NAO-type pattern are termed ‘Type-N’. We test the theory of Cohen and Saito (2003) by compositing winters into ‘Type-A’ and ‘Type-N’ and find that October Eurasian snow cover performs better in ‘Type-A’ than in ‘Type-N’ years. However, the number of ‘Type-A’ winters (1961-2000) is just ten and their cross-validated NAO_{DJF} hindcast skill is not statistically significant ($r_s = 0.63$, $p = 0.06$). Therefore, October Eurasian snow cover has predictive potential in only $\sim 25\%$ of winters. Furthermore, there is currently no *a priori* method for determining whether a winter will evolve into ‘Type-A’ or ‘Type-N’. We conclude that Eurasian snow cover may provide predictability for the winter AO but summer NH snow cover offers greatest predictability for the NAO_{DJF} .

3.4.4 Physical basis for summer ΔT influence on upcoming winter NAO_{DJF}

Extratropical seasonal predictability is normally assumed to be low, due to the dominating influence of chaotic weather systems, particularly in the autumn and winter seasons. This appears to contradict the links we observe between summer subpolar air temperature gradients, snow cover and the upcoming winter climate. There is little persistence intrinsic to the extratropical atmosphere and the decorrelation time is known to be ~ 10 days. Therefore, for the summer/winter link to

Period	Lagged Predictor	NAO		May SST (SVD)				JJASO SST PC2				JJ $\Delta_x T_{60N-70N}$			
		$r[\text{lag-1}]$	(p)	r	(p)	$MSSS$	(p)	r	(p)	$MSSS$	(p)	r	(p)	$MSSS$	(p)
1900-2001	CRU NAO _{DJF}	0.13	(0.32)	0.31	(<0.01)	9	(<0.01)	0.14	(0.10)	0	(-)	0.27	(0.01)	7	(0.01)
	Hurrell NAO _{DJF}	0.24	(0.07)	0.32	(<0.01)	10	(<0.01)	0.26	(0.01)	6	(0.01)	0.31	(<0.01)	9	(<0.01)
	MSLP PC1 _{DJF}	0.20	(0.15)	0.42	(<0.01)	18	(<0.01)	0.15	(0.10)	1	(0.07)	0.25	(0.01)	6	(0.02)
1950-2001	CRU NAO _{DJF}	0.15	(0.45)	0.04	(0.30)	0	(-)	0.07	(0.31)	0	(-)	0.10	(0.19)	0	(-)
	Hurrell NAO _{DJF}	0.24	(0.21)	0.15	(0.14)	0	(-)	0.29	(0.02)	7	(0.02)	0.12	(0.19)	0	(-)
	MSLP PC1 _{DJF}	0.17	(0.43)	0.19	(0.12)	1	(0.16)	0.01	(0.46)	0	(-)	0.12	(0.17)	0	(-)

Table 3.5: As Table 3.3 except for lag-1 autocorrelation ($r[\text{lag-1}]$) of NAO_{DJF} indices and skill values from cross-validated hindcasts using the three NAO_{DJF} indices as predictors.

work, the memory of the summer atmospheric state must persist in a slowly-varying boundary variable that can feed back onto the atmosphere at a later time.

Following Saunders et al. (2003), we propose the following mechanism for the influence of summer climate on that of the upcoming winter: Summer (JJ) NH snow cover anomalies are negatively correlated with JJ subpolar near-surface air temperature (T2m) over northern Siberia and northwest Canada and positively correlated with JJ T2m over Southern Greenland. The resulting subpolar zonal temperature gradients induce a contemporaneous atmospheric circulation response and lagged SST response centred on the North Atlantic. The atmospheric response is characterised by anomalies in MSLP and midlatitude zonal wind. The SST response is characterised by basin-scale anomalies and the establishment of meridional gradients southeast of Newfoundland. When this pattern persists into winter, it could feed back onto the atmosphere to influence the sign and magnitude of the NAO_{DJF} .

There is good observational evidence to support this hypothesis. First, there is a significant link ($p < 0.01$) between changes in JJ Northern Hemisphere snow cover and JJ ΔT for the period 1972-2001 (Table 3.4). Table 3.3 shows that ΔT is the best overall NAO_{DJF} predictor and NH snow cover may be the largest contributor to ΔT . Second, the atmosphere adjacent to the ΔT regions is significantly perturbed during June and July. Figure 3.5 shows zonally averaged zonal wind anomalies as a function of height above the ΔT regions and the North Atlantic (1950-2001) before and during high minus low JJ ΔT tercile years. Prior to June and July the only consistent signal for all three regions is located in the stratosphere at high latitudes during April. By May, there are significant anomalous easterlies over central Eurasia but not above North America or the North Atlantic. In June and July, a significant signal occurs over Eurasia and North America characteristic of a strengthened polar vortex and a weakened midlatitude jet. A corresponding teleconnected signal is seen over the North Atlantic in June and July which extends from the surface to the lower stratosphere. A band of significant anomalous westerlies is centred between 50°-60°N, with anomalous easterlies to the North and South. This highlights that during extreme ΔT years, contemporaneous subpolar atmospheric teleconnections are formed between northern Eurasia, northern Canada and the North Atlantic.

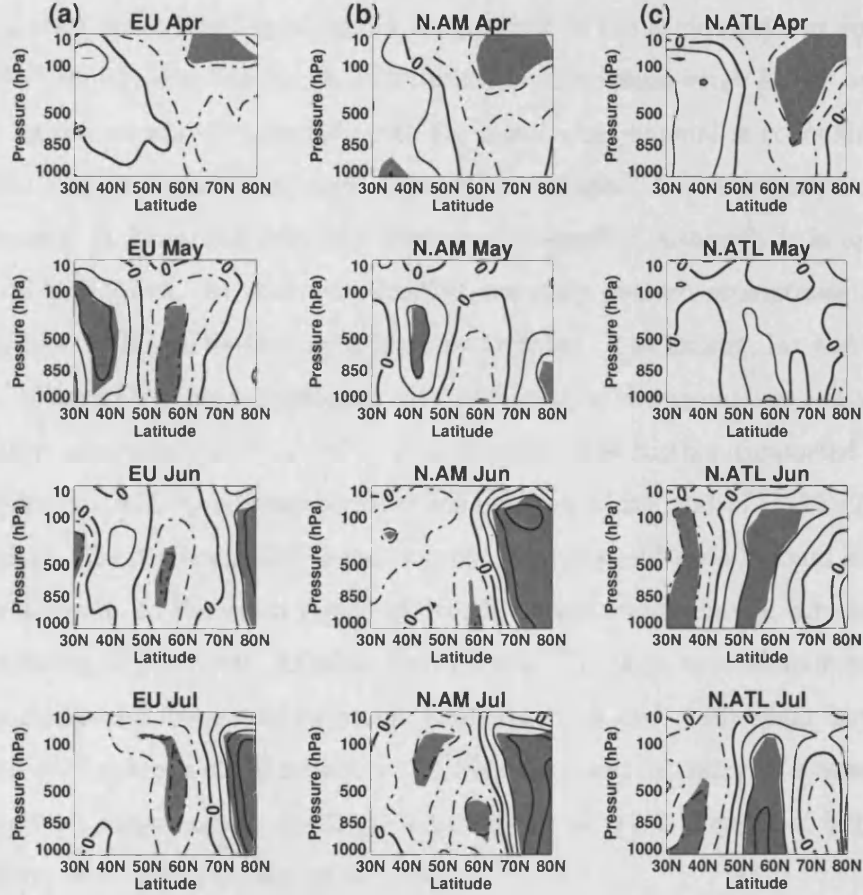


Figure 3.5: Composite vertical cross section of zonally averaged zonal wind anomalies based on high minus low terciles of June-July ΔT index 1950-2001. Zonal averages are calculated over (a) Eurasia, 25°E-70°E, (b) North America, 120°W-90°W and (c) North Atlantic, 50°W-20°W. Data are dimensionless standardised anomalies. Contour interval is 0.3 and dashed contours denote negative anomalies. Shaded areas denote values significant at less than 5% as determined by a Student's *t*-test.

Our proposed mechanism for the influence of summer climate on that of the upcoming winter is further supported by Figure 3.6, which shows the hemispheric-scale anomaly patterns in height-averaged (925 to 200 hPa) zonal wind, MSLP and SST in months before, during and following high minus low JJ ΔT tercile years 1950-2001. A clear strengthening of signals is apparent in the atmosphere in June and in the SST in July over the North Atlantic sector. The zonal wind, MSLP and SST signals appear causally connected, with the zonal wind anomalies collocated with gradients in MSLP anomalies and with SST anomalies. The atmospheric signal is prominent in June and July but dissipates thereafter, although it is apparent in the JJASO mean. In contrast, the SST anomaly pattern strengthens in July and persists, although weakening a little by October. The timing, lag and spatial pattern of the SST signal is consistent with its forcing by the anomalous atmospheric circulation associated with JJ ΔT . This conclusion is further supported by the significant ($p < 0.05$ for all time periods) link between JJ ΔT and JJASO SST (PC2) (Table 3.4). The persistent SST signal located southeast of Newfoundland in Figure 3.6 lies adjacent to the main region of North Atlantic cyclogenesis, which marks the beginning of the North Atlantic storm track. Through anomalous meridional SST gradients and associated turbulent heat, moisture and momentum fluxes, the observed SST pattern could influence the frequency and intensity of extratropical cyclones and, subsequently, NAO_{DJF} (Rodwell et al. 1999, Peng and Whittaker 1999, Peng et al. 2003, Cassou et al. 2004).

Figure 3.7 shows the mean anomalies in vertically averaged zonal wind, MSLP and SST for the winter seasons before and following high minus low JJ ΔT terciles. The situation preceding JJ ΔT shows a moderate positive NAO_{DJF} pattern in the zonal wind anomalies but SSTs close to climatology with no significant departures from the mean except near the equator. This further underlines that NAO_{DJF} predictability from JJ ΔT does not derive from interannual NAO_{DJF} persistence. The winter SST pattern after extreme JJ ΔT episodes exhibits a strong meridional gradient off Newfoundland, which has persisted from the previous summer.

Next, we attempt to show whether the evolving anomalies seen in Figure 3.6 indicate a real response to JJ ΔT or merely represent increased climatological vari-

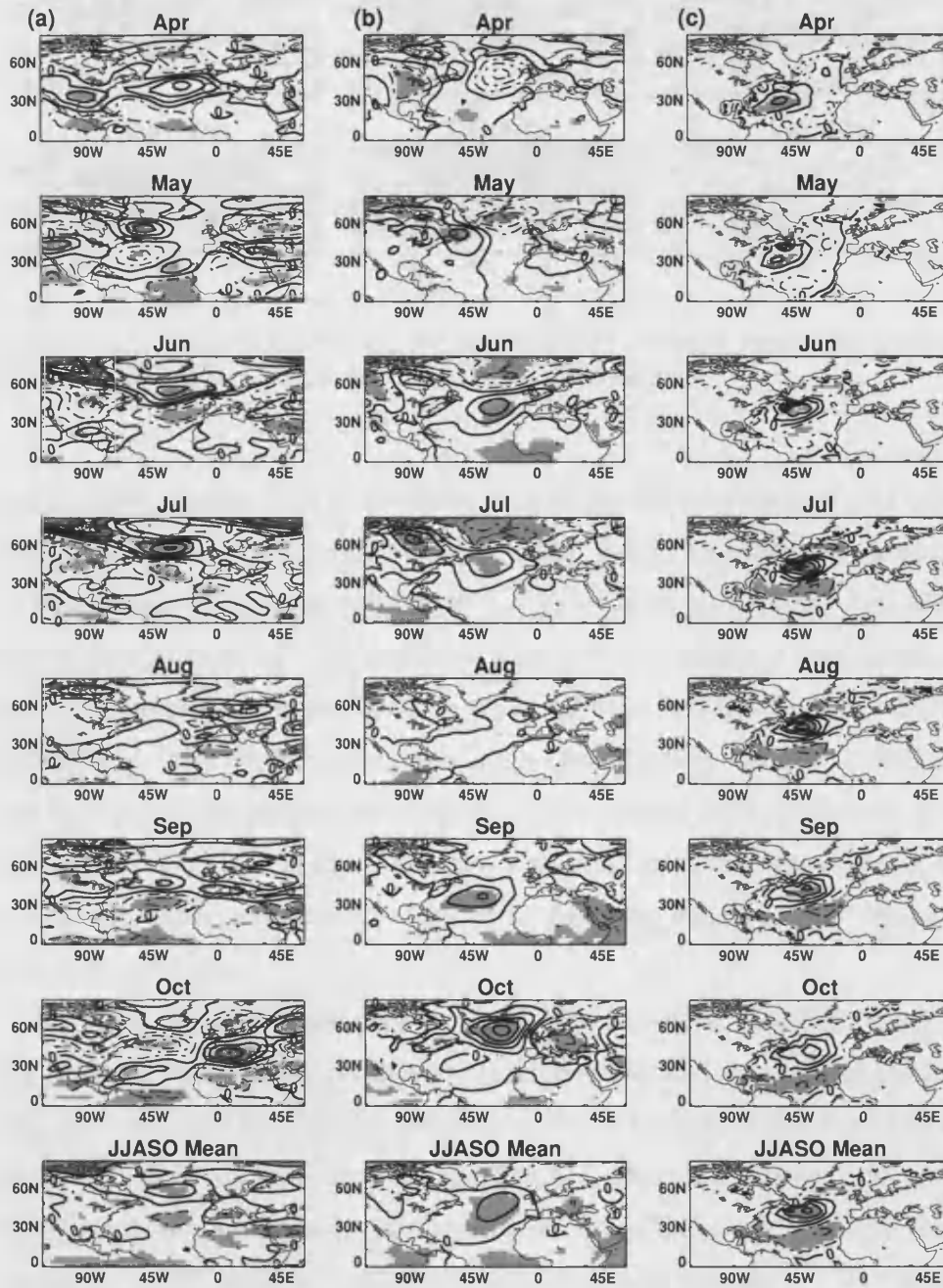


Figure 3.6: Composite of (a) vertically averaged (925-200 hPa) North Atlantic sector zonal wind anomalies, (b) North Atlantic sector mean sea-level pressure anomalies (MSLP) and (c) North Atlantic SST anomalies based on high minus low terciles of June-July ΔT index. Contour interval is (a) 1 m s^{-1} , (b) 1 hPa and (c) 0.2°C . Dashed contours denote negative anomalies and zero contour is labelled. Shaded areas denote values significant at less than 5% as determined by a Student's *t*-test.

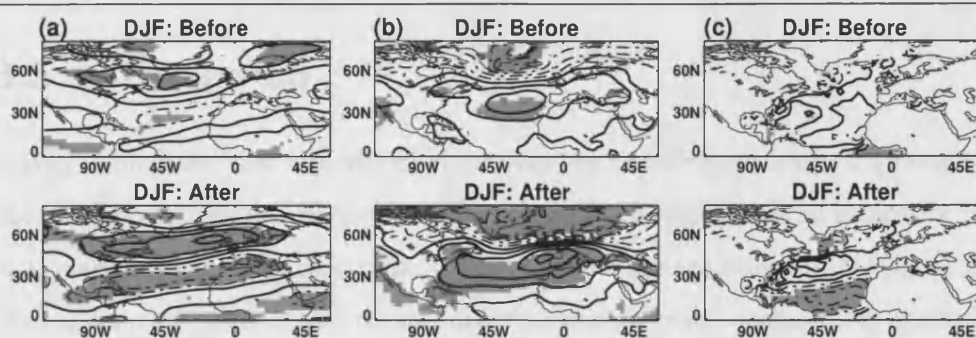


Figure 3.7: As Figure 3.6 except for winter (DJF) seasonal means (a) preceding and (b) following high minus low June-July ΔT terciles.

ance in those months. This is quantified by plotting the percentage of grid cells in each month that are locally significant ($p < 0.05$). Figure 3.8 shows the percentages for zonal wind over the region $[30^{\circ}\text{N}-80^{\circ}\text{N}, 120^{\circ}\text{W}-40^{\circ}\text{E}]$ and for SST over the region $[0^{\circ}-65^{\circ}\text{N}, 100^{\circ}\text{W}-0^{\circ}]$. The horizontal lines denote estimates of field significance ($p < 0.05$) assuming 30 spatial DOF in the zonal wind field and 15 spatial DOF in the SST field. These values are appropriate because Livezey and Chen (1983) state that hemispheric atmospheric fields contain ~ 30 -60 spatial DOF. The actual spatial DOF may oscillate around these estimates depending on the month. However, even with a conservative estimate for DOF of 15 the zonal wind and SST signals are both field significant.

The zonal wind signal peaks in June and July, consistent with Figure 3.6. The SST response develops one-month later, peaking in July and August, and persisting until October. By October, the majority of the SST signal comes from the subtropics. The SST response decreases to December, which is associated with greater variability in the atmospheric pattern August through October. Neither the atmospheric or SST patterns exhibit any winter-to-winter persistence. We therefore conclude that the anomalous atmospheric circulation associated with JJ ΔT is leading North Atlantic SST variability with a time lag of ~ 1 month.

3.5 Discussion

NAO_{DJF} hindcast skill is nonstationary over the twentieth century, implying that the predictive relationships vary in time. There are several potential causes for nonstationarity. First, data quality may contribute to the variations in NAO_{DJF} skill. During the early part of the record, observations are more sparse than during the recent 1972-2001 period of highest skill. This reduction in data quality may degrade NAO_{DJF} hindcast skill and contribute to nonstationarity. Second, nonstationarity in predictive relationships may arise from low-frequency multiannual or decadal oscillations in the predictor and/or predictand indices. An analysis of spectral coherency (Bloomfield 2000) was performed on the predictors and NAO_{DJF} indices to determine the dominant time scales of interaction. The results (not displayed) show that the preferred oscillatory period for each predictor with the NAO_{DJF} differs depending on the assessment period. For example, the MJJAS ΔT index and CRU NAO_{DJF} index are significantly coherent at a period of 7 years 1900-2001, 4 years 1950-2001 and 8 years 1972-2001. These changing multiannual periods for spectral coherency may contribute to the observed nonstationarity in predictive relationships. A further cause of the nonstationarity may be competing influences from SST forcings outside the North Atlantic, in particular, variations in the strength of ENSO (Sutton and Hodson 2003).

Using our standardised hindcast procedure, the May SST (SVD) predictor proposed by Rodwell and Folland (2002) exhibits no significant NAO_{DJF} skill. Differences between our methodology and the original methodology include the assessment time periods, NAO_{DJF} indices, skill measures, block elimination and sea-ice removal. These differences may explain a proportion of the skill. A simple spatial correlation analysis illustrates why May SST may offer less NAO_{DJF} predictability than JJASO SST. Correlating North Atlantic SST and the upcoming CRU NAO_{DJF}, the region off Newfoundland exhibits an area of significant ($p < 0.05$) correlation three times as large with JJASO SST than with May SST (not shown). The zonal wind composite for May in Figure 3.6 also shows dipolar anomalies overlying this region. This implies that May SST could be related to NAO_{DJF} through

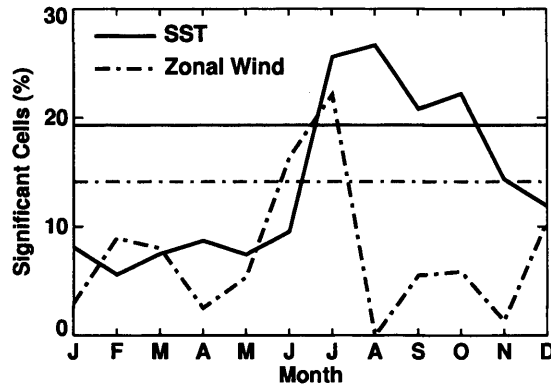


Figure 3.8: Area average of absolute standardised monthly mean anomalies of SST (solid line, dimensionless) and vertically averaged (925-200 hPa) zonal wind (dashed line, dimensionless) associated with high minus low terciles of June-July ΔT index. Values represent area average over the region $[20^{\circ}\text{N}-65^{\circ}\text{N}, 100^{\circ}\text{W}-0^{\circ}\text{E}]$ for SST and the region $[40^{\circ}\text{N}-65^{\circ}\text{N}, 120^{\circ}\text{W}-40^{\circ}\text{E}]$ for zonal wind.

our proposed ΔT mechanism but the predictive link from May is weaker. This concurs with the findings of Radcliffe and Murray (1970), who show that autumn SSTs off Newfoundland offer greatest predictability for the subsequent winter.

SST persistence associated with JJ ΔT is shown to occur from summer into late autumn. This suggests that the JJ ΔT link to winter climate is independent of interannual NAO_{DJF} persistence associated with North Atlantic SSTs submerged beneath the mixed layer during summer (Deser et al. 2003). The decrease in the SST signal from October to December seen in Figure 3.8 indicates that the ΔT -induced SST anomalies are not significant departures from climatology in these months. However, by November and December the NAO circulation dominates over the North Atlantic sector and the passive SST response will be NAO-driven more than ΔT -driven. Therefore, the role of early summer ΔT is only to enhance summer and autumn SSTs that can feed back onto the atmosphere prior to the dominant winter NAO circulation becoming established.

Summer NH snow cover is assumed to play a significant role in establishing ΔT (see Table 3.4). However, the JJ ΔT index produces higher NAO_{DJF} skill than does JJ NH snow cover (see Table 3.3), which implies that ΔT is the more important NAO_{DJF} predictor. Since JJ NH snow cover explains $\sim 50\%$ of the variance in

NCEP/NCAR JJ ΔT , other climate forcings must explain the remaining 50%. Another possibility is that the ΔT index forces changes in NH snow cover and JJ NH snow is not critical to the physical mechanism. In this case, an as yet unidentified third variable could influence ΔT . Further investigation is required to deduce the exact order of these relationships.

The mechanism linking summer ΔT to NAO_{DJF} relies on the persistence of summer/autumn SST anomalies and their feedback onto the NAO_{DJF} . However, ΔT produces better NAO_{DJF} hindcast skill than JJASO SST (PC2) despite the shorter time lag between this SST signal and the NAO_{DJF} . Several possible explanations exist for this apparent contradiction. First, the JJASO SST (PC2) predictor may not capture completely the feedbacks either from JJ ΔT or to the NAO_{DJF} . This explanation would be consistent with the use by Saunders and Qian (2002) of two JJASO lagged modes (EOF 2 and 5) of North Atlantic SST variability to enhance NAO_{DJF} skill. Second, the feedback from ΔT to the NAO_{DJF} may occur through a third variable and the associated JJASO SST pattern is a side-effect of the anomalous summer ΔT atmospheric circulation. Third, summer ΔT could be physically unrelated to NAO_{DJF} and they correlate through random interaction of their low-frequency oscillations. Our results support a physical relationship between summer and winter climate but further investigation using coupled dynamical models is required to fully understand this link.

Our findings suggest that it is the NH subpolar regions rather than the mid-latitudes or the tropics which provide the best extended-range NAO_{DJF} seasonal predictability. This contrasts with recent thinking based on atmospheric GCM experiments, which indicate that variations in tropical SSTs are of primary importance for explaining the NAO_{DJF} trend 1950-2000 (Hurrell et al. 2004). While our analysis focuses solely on interannual variability by using detrended time series throughout, we find that using data with trends included also gives similar results.

3.6 Summary and conclusions

This Chapter presents a detailed assessment of the current levels of seasonal empirical predictability of the North Atlantic Oscillation (NAO). A standardised hind-cast procedure is used to validate four previously published lagged predictors of the NAO_{DJF} . A new predictor based on the zonal gradient in summer Northern Hemisphere subpolar 2 m air temperature (ΔT) is also examined. Over three extended assessment periods out to 100-years, summer ΔT is most skilful in predicting NAO_{DJF} . For the period 1900-2001, May-September mean ΔT offers highest skill (~ 6 -9% improvement over climatology). Since 1972, June-July mean ΔT produces highest skill (~ 35 % improvement over climatology). The increase in NAO_{DJF} predictability since 1972 is observed for all predictors except those derived using SSTs and coincides with a period of high decadal NAO_{DJF} variability. The intervals of best predictability from the $\Delta T/NAO_{DJF}$ link coincide with positive trends in the NAO_{DJF} index in the early and late twentieth century. However, ΔT performs equally well predicting above or below median NAO_{DJF} seasons.

Evidence is presented supporting a physical link between summer ΔT and NAO_{DJF} . First, there is a strong contemporaneous association between summer ΔT and Northern Hemisphere snow cover extent. Snow cover plays a significant role in establishing ΔT in summer. During subsequent months, the atmospheric response to ΔT is centred on the midlatitude North Atlantic. Circulation anomalies over the ocean lead by ~ 1 month a pattern of North Atlantic SST, which persists into autumn. SST persistence is strong off southeast Newfoundland, which coincides with the main region of North Atlantic cyclogenesis. Therefore, through meridional SST gradients and in situ surface fluxes, the SST pattern could influence the formation of extratropical cyclones and, subsequently, the NAO_{DJF} .

Chapter 4

Coupled GCM Representation of Observed Snow Cover and Associated Atmospheric Variability 1972-2002

4.1 Introduction

Summer Northern Hemisphere (NH) snow cover was identified in Chapter 3 as a skilful lagged predictor of the winter NAO (NAO_{DJF}). However, observations of NH snow cover prior to 1972 are considered unreliable for climatological investigations (Robinson et al. 1993). The limited availability of observational data leads to problems in assessing stationarity and statistical significance of these snow/atmosphere links. Therefore, general circulation model (GCM) simulations are employed to investigate these links in more detail. Atmospheric GCM experiments have been conducted to simulate the observed (Cohen and Entekhabi 1999) relationship between autumn season Eurasian snow cover and the upcoming NH winter climate (e.g., Gong et al. 2003). To date, no GCM experiments have been conducted to simulate the observed (Saunders et al. 2003) link between summer NH snow cover and the NAO_{DJF} . Furthermore, the use of coupled GCM (CGCM) integrations in previous snow/atmosphere investigations is very limited. In this Chapter and the next, we address the research questions posed in Section 1.5.4. The main question is “How well does a CGCM represent the observed link between summer NH snow cover and the NAO_{DJF} ?”. We investigate this problem using the output from a twentieth century climate integration run on a leading CGCM.

This study is divided into two parts. The first part, presented in this Chapter,

compares the observed temporal and spatial variability in snow cover for all seasons to that simulated in the CGCM. Previous validation studies of snow cover in GCM simulations have been performed over relatively short periods (e.g., 1979-1995 in Frei et al. (2003)). CGCM and observational data are compared for the 31-year period 1972-2002. The nature of the CGCM's seasonal *in situ* atmospheric response to snow cover over Eurasia and North America is examined to establish how well the CGCM simulates the observed interaction between snow cover and the atmosphere.

The second part, presented in Chapter 5, focuses on the observed link between summer NH snow cover and the NAO_{DJF} . First, a comparison is performed of observed and CGCM NAO_{DJF} variability 1904-2002. Second, NAO_{DJF} predictability from lagged snow cover in the CGCM is quantified using lagged correlation and cross-validated NAO_{DJF} hindcasts. Third, we examine whether the physical mechanisms linking summer snow cover and the NAO_{DJF} in the CGCM are the same as in observations. Fourth, we investigate the stationarity in NAO_{DJF} predictability from lagged snow cover is quantified using the CGCM data for the extended period 1904-2002. We conclude this study with a review of the limitations in using a CGCM for snow/atmosphere investigations. A number of recommendations for potential improvements to CGCM formulations are made.

This Chapter is structured as follows. Section 4.2 provides a brief description of the CGCM used in this study. Details are also presented of the twentieth century climate integration from which the CGCM data are taken. The correlation and wavelet analysis methods employed in this study are discussed in Section 4.3. Section 4.4 describes the conversion of CGCM snow output to an index of snow covered area (SCA). Here, the comparison between the observed and CGCM representation of the temporal and spatial variability of SCA is presented. Section 4.5 examines the *in situ* atmospheric temperature and surface flux responses to regional SCA. A discussion of the findings is presented in Section 4.6 and a brief summary is given in Section 4.7.

4.2 ECHAM4/OPYC3

The CGCM scheme selected for use in this study is the ECHAM4 atmospheric GCM coupled to the OPYC3 oceanic GCM, which was jointly sponsored by the Max Planck Institut für Meteorologie (MPI) and the Deutsches Klimarechenzentrum (DKRZ). This modelling scheme was selected for three main reasons. First, the ease of access to a comprehensive set of output data from a long (1860-2050) historical climate integration included in CMIP. These data were provided through the Hamburg World Data Center for Climate CERA Internet data portal. Second, to date, only a small number of published studies have used the ECHAM4/OPYC3 CMIP data to perform comparisons with observations (e.g., Hanssen-Bauer and Forland 2001, Skaugen and Tveito 2004). This work is the first to validate ECHAM4/OPYC3 snow cover representation and therefore makes effective use of this large repository of CGCM data. Third, successful results have been achieved in investigations into snow/atmosphere interaction using the earlier ECHAM3 atmospheric GCM (e.g., Gong et al. 2003). The present study therefore builds on previous investigations and examines how coupling ECHAM4 to the ocean affects the representation of snow/atmosphere interaction.

GCM formulation

The ECHAM4 atmospheric GCM has evolved from the spectral numerical weather forecasting model of the European Centre for Medium-range Weather Forecasts (ECMWF) and has been modified extensively at MPI for climate applications (Chen and Roeckner 1996). Compared to ECHAM3, ECHAM4 includes improvements to cloud and radiation, moisture transport, convection, boundary-layer turbulence and land surface schemes. A detailed description of the dynamical and physical structure (and the simulated climatology) of ECHAM4 is documented in Roeckner et al. (1996). The atmosphere has 19 hybrid vertical levels and the uppermost level is centred at 10 hPa. Both annual and diurnal cycles are included. The time-step employed in all integrations is 24 minutes with radiation fluxes updated every 2 hours (Chen and Roeckner 1996). Turbulent surface fluxes are calculated from

Monin-Obukhov similarity theory. Within and above the atmospheric boundary layer, a higher-order closure scheme is used to compute the turbulent transfer of momentum, heat, moisture and cloud water (CMIP 2005).

Snow is a prognostic variable in ECHAM4 and is augmented by snowfall and depleted through sublimation and melting. The snow pack core temperature is computed using the heat diffusivity/capacity of ice in regions of permanent continental ice and for bare soil where snow depth is less than 2.5 cm. When snow depth is greater than 2.5 cm, the snow pack core temperature is derived using an auxiliary conduction method. The temperature at the surface of the snow pack is extrapolated from the core temperature and is constrained not to exceed the snow melt temperature (0°C). The soil moisture reservoir is augmented by melting snow. Ice and snow albedo are temperature dependent and the latter is also a function of the fractional forest area in a grid box. The ECHAM4 snow output field is gridded snow depth in metres, where each grid cell is assumed to have uniform snow depth.

A new global data set of land surface parameters is used in ECHAM4 based on Claussen et al. (1994). The parameters are constructed from the the major 'Olson' ecosystem complex. A land soil model is used that comprises the budgets of heat and water in the soil, the snow pack over land and the budget of land ice. Vegetation effects such as the interception of rain and snow by the canopy are parameterised in a highly idealised way (CMIP 2005). The atmosphere is coupled to the OPYC3 ocean GCM, whose name derives from its isopycnal vertical coordinate system. Full details of OPYC3 are provided by Oberhuber (1993). ECHAM4 and OPYC3 are coupled quasi-synchronously and exchange information once a day (Roeckner et al. 1999).

GSDIO integration

In this study we employ data from the 1860-2050 coupled sulphate aerosol, greenhouse gas and ozone integration (GSDIO) performed using ECHAM4/OPYC3. GSDIO was a transient integration that included observed historical greenhouse gas, direct and indirect sulphate aerosol forcing 1860-1990. After 1990, greenhouse gas

and aerosol emissions are constrained to follow the IPCC emissions scenario IS92a. This scenario prescribes effective (representing all greenhouse gases) CO₂ emissions increasing at 1% per year.

The spin up process for the coupled ECHAM4/OPYC3 run is used to calculate the flux adjustments required to prevent the GCM climate from ‘drifting’ during transient simulations. First, ECHAM4 was integrated for 20 years with prescribed SSTs and sea ice extents until quasi-equilibrium was reached. Second, OPYC3 was spun up for 500 years forced with climatological surface fluxes from the 20-year uncoupled ECHAM4 run. Third, OPYC3 was run for a further 500 years forced by daily anomalies of heat and freshwater fluxes derived from the 20-year uncoupled ECHAM4 run. Fourth, ECHAM4 and OPYC3 were coupled and integrated for 100 years, while restoring the SST and salinity towards climatology. During this coupled run, annual mean flux adjustments for heat and freshwater were derived and frozen after 100 years. Fifth, the frozen flux adjustments were applied throughout the subsequent GSDIO integration.

The GSDIO integration was performed using a spectral T42 resolution (Cartesian resolution $\sim 2.8^\circ \times 2.8^\circ$). OPYC3 was run with 11 ocean vertical density levels at the same T42 resolution as ECHAM4. To facilitate the comparison of the GSDIO fields with gridded observational data, all GSDIO fields were interpolated from the T42 spectral grid to a regular $2.5^\circ \times 2.5^\circ$ lat-lon grid. This was achieved using a method of bilinear interpolation. Although increasing the apparent spatial resolution, this method is likely to result in a loss of some information where grid cells overlap. However, as the difference between the mean T42 and 2.5° grid box sizes is relatively small, the loss of information was found not to affect our results. Monthly mean anomalies are employed and seasonal averages are computed from the monthly anomalies. The ECHAM4/OPYC3 GSDIO coupled integration will, henceforth, be referred to as ‘ECHAM4’.

4.3 Methodology

4.3.1 Correlation

Relationships between two time series are assessed using the Spearman rank correlation coefficient (r_{RANK}) instead of the traditional Pearson correlation coefficient (r). Rank correlation does not suffer from the problems associated with the Pearson correlation, namely, sensitivity to outliers and underestimation of nonlinear effects (Wilks 1995). The Pearson coefficient is used in this study only for computation of pattern correlations and autocorrelations. The Spearman rank correlation takes the form

$$r_{RANK} = 1 - \frac{6 \sum_{i=1}^N D_i^2}{N(N^2 - 1)}, \quad (4.1)$$

where D_i is the difference in rank between the i th pair of data values and N is the length of the series. Where there are tied values, each is assigned the average rank of the group.

Statistical significance is assessed through a transformation of the correlation coefficient. The r_{RANK} coefficient for a sample size $N > 20$ follows an approximately normal distribution with zero mean and $\sigma = 1/\sqrt{(N-1)}$ (Wilks 1995). The transformed variable is

$$z = r_{RANK} \sqrt{(N-1)} \quad (4.2)$$

Each correlation is tested against H_0 that the time series are uncorrelated ($z = 0$). As in Chapter 3, the number of degrees of freedom is reduced by the factor τ (see Section 3.2.3).

4.3.2 Low-frequency variability

Wavelet analysis

Wavelet analysis or wavelet transformation (WT) is a tool for analysing local variations in spectral power (Torrence and Compo 1998). WT employs a family of base functions (termed ‘wavelets’) that are scaled and translated to map out the entire

time-frequency domain. A criticism of wavelet analysis is that the choice of wavelet is somewhat subjective. However as noted by Torrence and Compo (1998) the same is true for traditional transforms such as Fourier, Bessel and Legendre. A physical signal should be independent of the choice of wavelet. Lau and Weng (1995) recommend that the analysing wavelet should bear some resemblance to the signal of interest. Since climate signals can be expected to be smooth and continuous, this leads naturally to the choice of non-orthogonal complex bases such as the Morlet and Paul functions (for a full discussion see Lau and Weng (1995)). The Morlet waveform is the most common and is selected for this analysis.

WT requires the data to be cyclic, which means that the beginning of the data is assumed to map smoothly onto the end. In general this is not the case and discontinuities are present at the endpoints which introduce spurious harmonic components. One way of reducing this problem is to pad the ends of the time series with zeros prior to transformation and remove them after the transformation. While this reduces edge effects, they are not eliminated entirely and as one goes to larger scales the amplitude of the spectrum near the ends is reduced as more zeros enter the analysis. The ‘cone of influence’ (COI) is the region in the wavelet spectrum in which edge effects become important.

The COI is defined as the e -folding time for the autocorrelation of wavelet power at each scale. The wavelet power of a discontinuity at the edge drops by a factor of e^{-2} over this period, therefore edge effects will be negligible beyond this point (Torrence and Compo 1998). The size of the COI at each time period provides information on the decorrelation time for a peak in the power spectrum. Peaks which last longer than this are likely to represent true harmonic components and not random fluctuations.

Peaks in the wavelet power spectrum are tested against a null hypothesis that the time series under investigation has the spectral characteristics of red noise. Therefore, it is assumed that the time series can be modelled as a lag-1, autoregressive process

$$x_i = \alpha x_{i-1} + z_i \quad (4.3)$$

where α is the lag-1 autocorrelation coefficient and the z_i are Gaussian white noise.

Autocorrelation

The prefix *auto*- in autocorrelation denotes the correlation of a variable with itself.

The Pearson autocorrelation of a time series \mathbf{X} at lag k is

$$\rho_X(k) = \frac{1}{\sigma^2} \text{Cov}(\mathbf{X}_i, \mathbf{X}_{i+k}). \quad (4.4)$$

Together, the collection of autocorrelations computed for various lags are termed the autocorrelation function (Wilks 1995). This is displayed graphically, with the autocorrelations plotted as a function of lag. Pearson correlation is used as it is more traditionally employed to compute autocorrelation than Spearman rank correlation.

4.4 Comparison of observed and ECHAM4 snow cover variability

4.4.1 Derivation of ECHAM4 snow cover indices

As discussed in Section 1.5.2, it is first necessary to create an index representing monthly mean CGCM snow covered area (SCA). This index can then be compared to the observed quantity, namely, SCA. We employ the method devised by Frei et al. (2003) because it is simple and compares closely to the method used for creating SCA indices from visible satellite images. The method converts gridded snow depth output to an index of fractional SCA by summing the areas of snow covered grid boxes. Fractional coverage, rather than absolute values in units of area, was preferred because data on different grids have different total land areas. Therefore, standardising by total land area provides a more robust comparison between the CGCM and observations.

A simple calculation is employed to determine what proportion of a grid cell with a certain snow depth is snow covered. Frei et al. (2003) estimated that SCA is a linear function of snow depth for grid cells with <2.5 cm snow depth (i.e., a

depth of 1.25 cm corresponds to a fractional SCA of 50% of the grid cell area). These authors summarized the conversion from snow depth to fractional SCA by the following equations:

$$f_i = 1 \quad d_i \geq d_c \quad (4.5)$$

$$f_i = d_i/d_c \quad d_i < d_c \quad (4.6)$$

$$f_R = \sum_R (f_i \times a_i), \quad (4.7)$$

where d_i =snow depth (m) for grid cell i ; d_c =critical snow depth (2.5×10^{-2} m); f_i =fractional snow coverage for grid cell i ; f_R =fractional coverage for region R ; and a_i =land area in grid cell i (m^2). ECHAM4 SCA indices are computed for the three land areas poleward of 20°N employed in the observational snow cover indices by Robinson et al. (1993). These areas represent Eurasia (EU), North America including Greenland (NA) and the entire Northern Hemisphere (NH), respectively.

In all months in the GSDIO simulation, Greenland was found to have zero snow depth lying on the ice sheet. Compared to observations, this introduced a negative bias to the ECHAM4 SCA indices because the land area of Greenland ($2.18 \times 10^6 \text{km}^2$) was omitted. To remove this bias in the SCA indices, the land area of Greenland is included as completely snow covered in all months of the year for the NA and NH indices. This corresponds closely to the snow coverage seen over Greenland in observations. Greenland is classified as a glacier in ECHAM4 and analysis of the surface albedo (α) showed that Greenland is prescribed with $\alpha \approx 0.8$ in all months during the GSDIO run. The potential limitations in the atmospheric response to a lack of Greenland snow cover must be considered when interpreting the ECHAM4 results. Henceforth, all snow cover data used have been bias-corrected over Greenland.

4.4.2 Spatial variability

The simplest comparison of observed and modelled SCA is that of the respective seasonal mean spatial distributions. Grids of fractional SCA for ECHAM4 were produced and the seasonal means are shown in Figure 4.1 alongside observations 1972-2002. The seasonal differences between observations and ECHAM4 are also plotted. In all seasons the mean spatial extent in ECHAM4 compares closely with the observations. The greatest differences are seen during winter and autumn over the snow margins of southern Siberia and the central United States. These are regions where snow cover has its greatest interannual variability, particularly during autumn. The Tibetan plateau region shows large errors in all seasons but these are confined to a relatively small area. The Pearson pattern correlations (Wilks 1995) between observations and ECHAM4 are highest for winter, spring and autumn ($r_{DJF} = 0.88$, $r_{MAM} = 0.90$, $r_{SON} = 0.90$) and lowest during summer ($r_{JJA} = 0.82$). Overall, these results show that the annual cycle in SCA is well reproduced by ECHAM4.

We also compared observed and ECHAM4 spatial SCA extents using the monthly mean SCA over an entire region. Figure 4.2 shows the percentage difference in SCA between observations and ECHAM4 for NA, EU and the NH 1972-2002. Positive values denote months where ECHAM4 produces less snow than observed and negative values indicate the opposite. Over NA, ECHAM4 underpredicts SCA in all months except February, with largest errors in June and November. Over EU, ECHAM4 overpredicts February through May and underpredicts June through January. The NH pattern corresponds closely to the mean of the NA and EU plots. These results show that the largest magnitude errors in ECHAM4 are found in individual months rather than across one particular season. Specifically, the largest errors in SCA over the entire NH are seen in June, November and December. Frei et al. (2003) found the largest errors for the mean of 15 atmospheric GCMs occurred during spring and autumn over all regions.

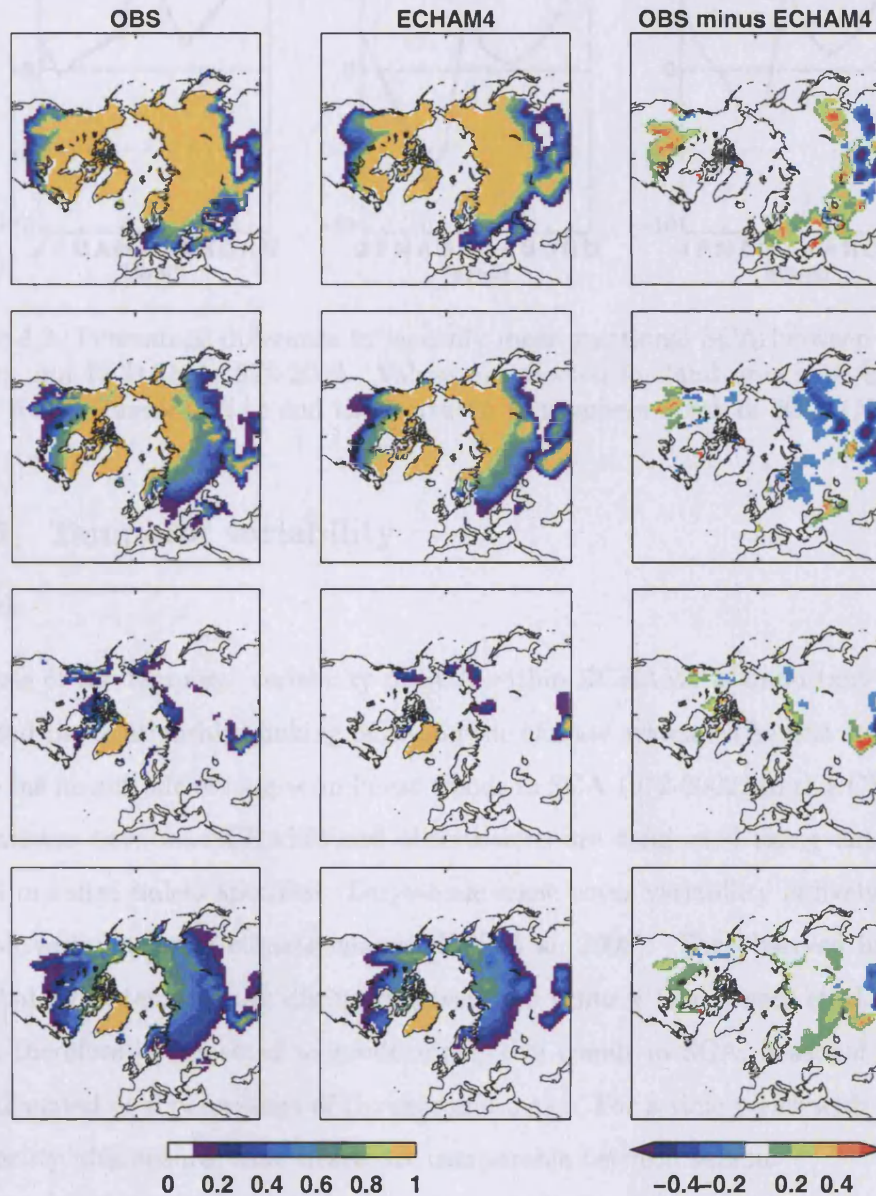


Figure 4.1: Spatial distribution of seasonal mean snow cover extent for winter (DJF), spring (MAM), summer (JJA) and autumn (SON) for observations (left), ECHAM4 (centre) and observations minus ECHAM4 (right). Colours indicate mean fractional coverage. Shading interval is 0.1, with values <0.1 not coloured.

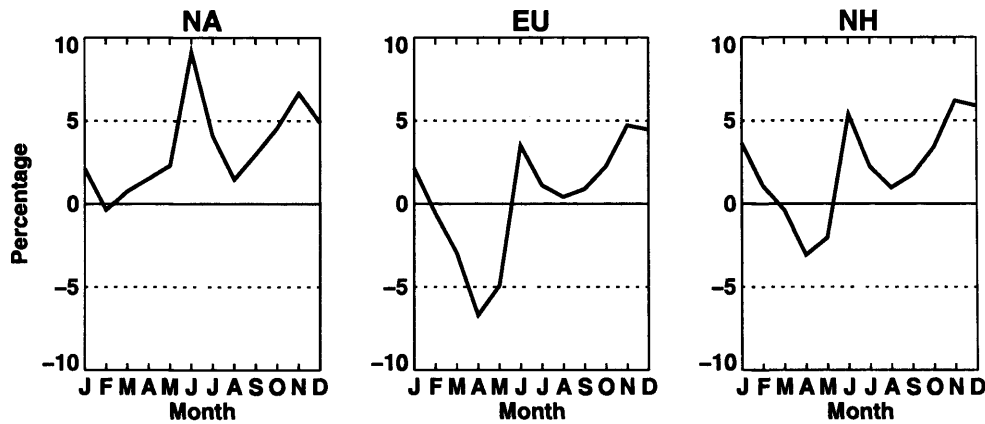


Figure 4.2: Percentage difference in monthly mean fractional SCA between observations and ECHAM4 1972-2002. Values are plotted for land area over Eurasia (EU), North America (NA) and the Northern Hemisphere north of 20°N (NH).

4.4.3 Temporal variability

Trends

Analysis of the temporal variability of SCA within ECHAM4 is important to understand the relationships linking SCA and the climate system. The first step is to assess the magnitude of long-term linear trends in SCA 1972-2002. In this Chapter, comparisons between ECHAM4 and observations are conducted using data with trends included unless specified. Large-scale snow cover variability is likely to be a sensitive indicator of climate change (Frei et al. 2003). The observed increase in global mean temperature during the twentieth century (e.g., Jones et al. 1999), would therefore be expected to produce negative trends in SCA. Seasonal trends are calculated as a percentage of the seasonal mean. For a time series with strong seasonality, this ensures that trends are comparable between seasons.

Table 4.1 shows the trends in observed and ECHAM4 SCA 1972-2002. In observations, largest trends for all regions are seen during spring and summer. The largest percentage change is -22.86% per decade for summer over EU. However, the mean spatial snow extent during summer is only 21% of that during spring, which means that similar snow amounts are disappearing due to negative trend during spring and summer. Smaller and non significant trends are found during winter

	Region	Annual	DJF	MAM	JJA	SON
OBS	NA	-2.02	-0.57	-2.10	-11.46	-0.18
	EU	-1.84	-0.15	-3.56	-22.86	1.08
	NH	-1.91	-0.31	-2.95	-15.30	0.53
ECHAM4	NA	-1.82	-0.77	-2.31	-1.97	-3.19
	EU	-3.96	-3.37	-3.64	-13.72	-5.73
	NH	-3.11	-2.42	-3.13	-5.34	-4.60

Table 4.1: Linear trends in annual and seasonal SCA indices 1972-2002 estimated using an ordinary least-squares regression fit. Trends are the percentage change in seasonal SCA per decade. Bold type denotes significant ($p < 0.05$) trends as determined by a Mann-Kendall test.

and autumn for all regions. ECHAM4 shows a different pattern, with significant trends seen in all seasons and in the annual mean over EU and NH but only in the annual mean over NA. The magnitude of trends over NA is uniformly lower than EU and NH during all seasons. The seasonal variation of the ECHAM4 trends also differs to observations. Values of $\sim -4\%$ per decade are found during all seasons in EU and NH, with the notable exception of summer in EU, which shows a much larger trend (-13.72%).

The negative trends in observed and ECHAM4 annual mean SCA 1972-2002 are in anticorrelation with the significant increase in near-surface air temperature over the same period. The ECHAM4 NH annual mean temperature 1972-2002 correlates strongly with observations ($r_{RANK} = 0.70$), which means that the temperature influence on SCA is approximately the same. Bamzai (2003) observed trends consistent with surface warming in the timing of both autumn snow onset and spring snow melt, particularly over EU. The negative annual mean trends have similar magnitudes in observations and ECHAM4. Seasonal trends are significant during spring in observations and ECHAM4 but during autumn are significant only in ECHAM4. ECHAM4 SCA appears to respond to increasing temperature with reduced seasonality compared to observations. Trends are stronger than observations during winter but weaker during summer, except over EU.

Interannual variability

The interannual correlation between GCM and observed SCA indicates the proportion of GCM SCA variability driven by boundary conditions and the proportion driven by internal atmospheric dynamics (Frei and Robinson 1998). The significant negative trends described above have an influence on year-to-year SCA variations. This influence is strongest during summer, when linear trend accounts for $\sim 20\%$ of the variance between observed and ECHAM4 SCA. Therefore, we present the comparison of interannual SCA variability using detrended data.

Figure 4.3 shows the interannual anomalies in observed and ECHAM4 detrended seasonal mean SCA for NA, EU and NH and the rank correlations between the indices 1972-2002. The overall correspondence between ECHAM4 and observations is very low. This shows little improvement on the accuracy achieved by the atmospheric GCMs investigated by Frei et al. (2003). All correlations are actually negative except those for summer, which peak for NH SCA ($r = 0.30$). The low correspondence could result from problems with the ECHAM4 internal dynamics and/or problems associated with the SCA response to boundary forcing. The transient nature of the CGCM integration means it is likely that the ECHAM4 and observed boundary forcings exhibit temporal and spatial differences 1972-2002. The interannual variability of ECHAM4 SCA is of similar magnitude to observations during spring. For the other seasons, variability is $\sim 20\%$ greater than observed during winter, ~ 3 times lower than observed in summer and around half observed during autumn.

The correlation between the NA, EU and NH SCA indices illustrates the relative importance of the EU and NA regions in the NH climate system. Table 4.2 shows r_{RANK} between the detrended regional SCA indices for annual and seasonal means 1972-2002. In general, NH SCA is correlated more strongly with EU than with NA and ECHAM4 correctly simulates this relative importance. Two important differences between observations and ECHAM4 are seen during summer. First, the links between NA/NH and EU/NH are highly significant ($r_{RANK} > 0.9$) in observations but are weaker in ECHAM4. Second, there is a significant ($r_{RANK} =$

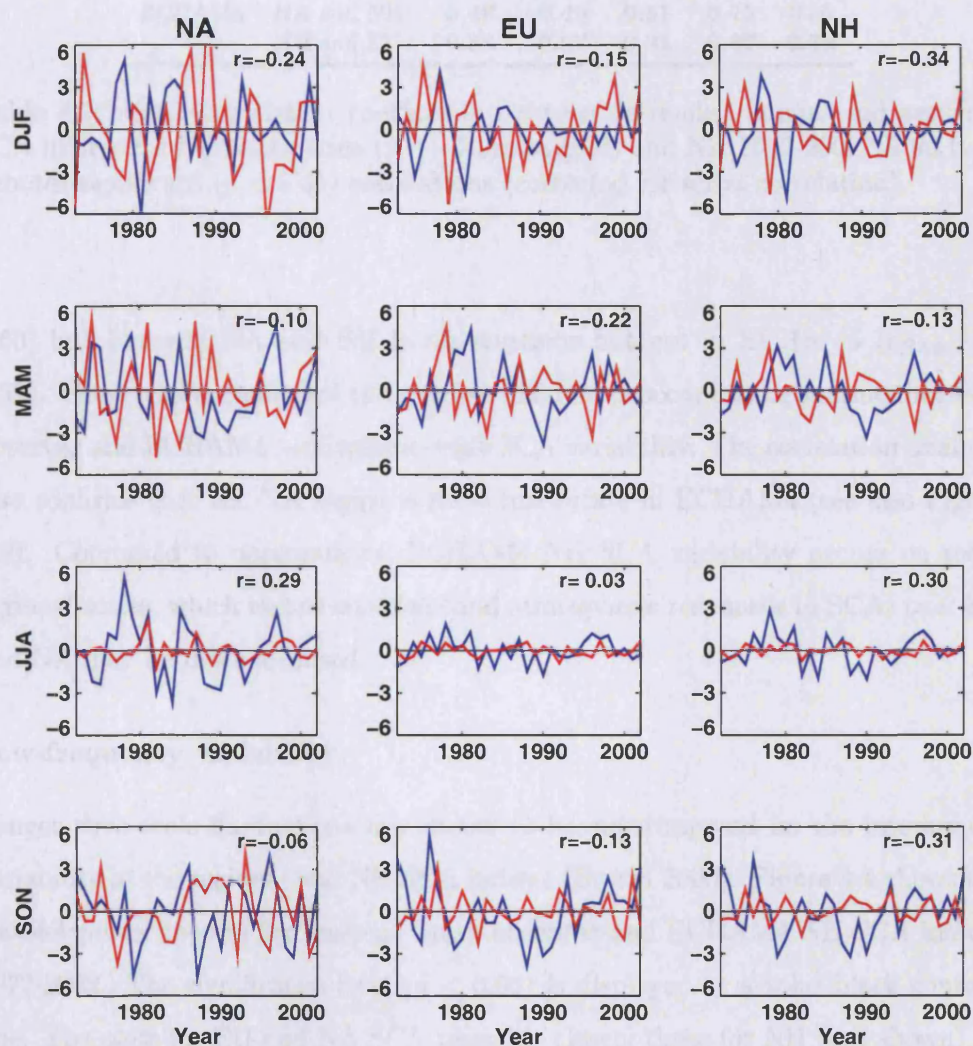


Figure 4.3: Time series of detrended seasonal mean SCA anomalies 1972-2002 for NA (left), EU (centre) and NH (right). Observed values are in blue and ECHAM4 values in red. Data are percentage SCA anomalies. Rank correlation coefficients are displayed.

	Regions	Annual	DJF	MAM	JJA	SON
OBS	NA and EU	0.30	0.26	0.07	0.66	0.33
	NA and NH	0.68	0.70	0.50	0.91	0.60
	NH and EU	0.88	0.82	0.85	0.90	0.94
ECHAM4	NA and EU	0.03	0.06	-0.03	0.39	-0.21
	NA and NH	0.48	0.49	0.51	0.75	0.36
	NH and EU	0.85	0.87	0.81	0.87	0.78

Table 4.2: Rank correlation coefficients between detrended annual and seasonal SCA indices for North America (NA), Eurasia (EU) and NH 1972-2002. Bold type denotes significant ($p < 0.05$) correlations (corrected for serial correlation).

0.66) link between NA and EU in observations but not in ECHAM4 ($r_{RANK} = 0.39$). These results show that the greatest differences occur during summer between observed and ECHAM4 hemispheric-scale SCA variability. The correlation analysis also confirms that the NA region is most inaccurate in ECHAM4 (see also Figure 4.2). Compared to observations, ECHAM4 NH SCA variability occurs on more regional scales, which means snowfall (and atmospheric responses to SCA) over EU and NA may be more localised.

Low-frequency variability

Longer time scale fluctuations are known to be superimposed on the interannual variability in the regional and NH SCA indices (Brown 2000). Figure 4.4 shows the wavelet power spectra for seasonal mean observed and ECHAM4 NH SCA indices 1972-2002. The significance level ($p < 0.05$) is displayed as a solid black contour line. The plots for EU and NA SCA resemble closely those for NH (not shown).

The observed SCA indices exhibit statistically significant variability only prior to 1990. This suggests that after 1990, the observed SCA spectrum has either been white or exhibited variability at periods greater than 16 years, which lie outside the COI with 30 years of data. During winter, the dominant period before 1990 is ~ 7 years, while for spring and summer it is ~ 2 -3 years. Only during autumn is strong power seen at decadal time scales, although this is not significant and lies outside of the COI. The ECHAM4 plots show little significant power at any period during

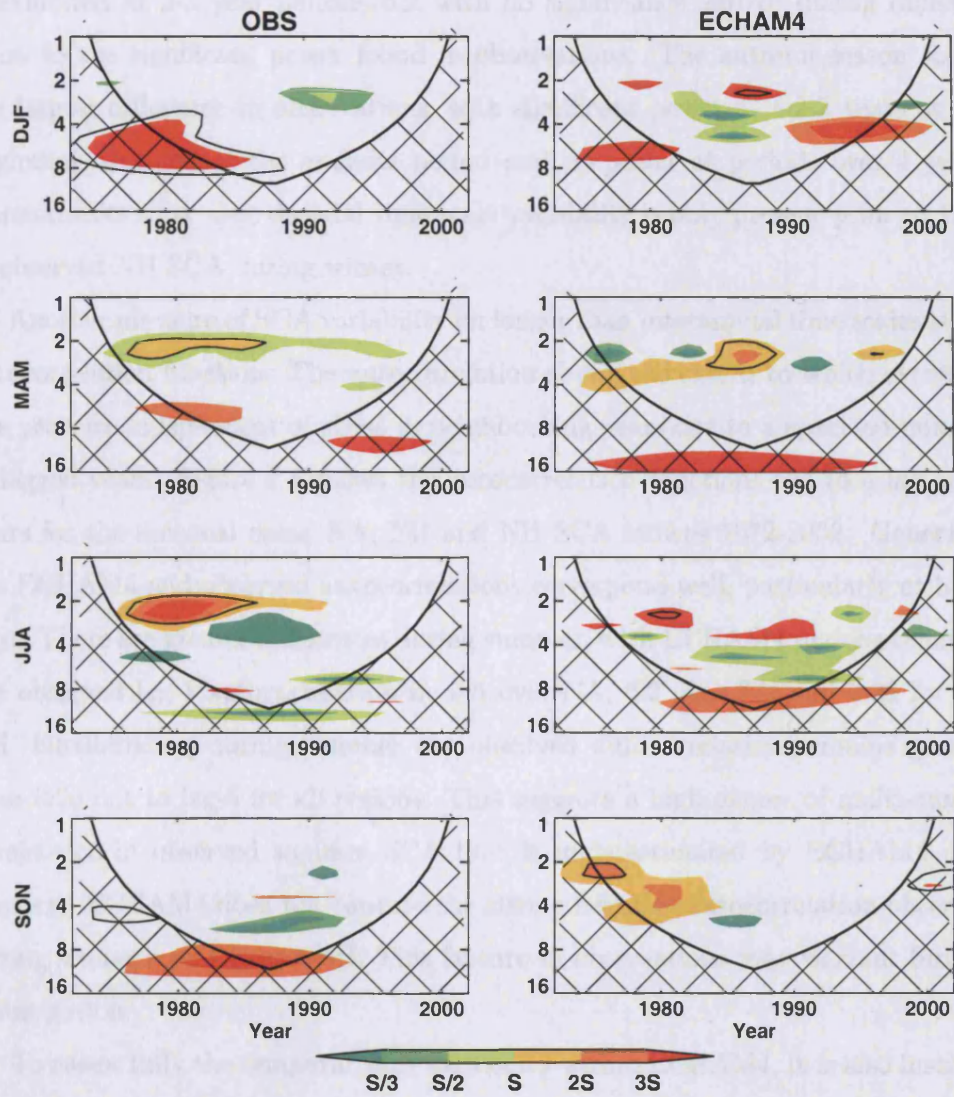


Figure 4.4: Wavelet power spectra of seasonal mean NH SCA 1972-2002 for observations (left) and ECHAM4 (right). y -axis shows period in years and cone of influence is indicated by cross-hatching. Colour shading indicates wavelet power at levels $S/3$, $S/2$, S , $2S$ and $3S$, where S is the standard deviation of the seasonal wavelet power. Black contour denotes significant ($p < 0.05$) power.

any season. During winter, there is an area of strong power at ~ 7 years around 1980 but this is not significant. Furthermore, during spring and summer power is exhibited at 2-3 year periods but with no significance and/or during different years to the significant power found in observations. The autumn season shows the largest difference to observations, with significant power at $\sim 2-3$ years at the beginning and end of the analysis period and no power at periods over 4 years. These results show that decadal time scale variability is only present prior to 1990 in observed NH SCA during winter.

Another measure of SCA variability on longer than interannual time scales is the autocorrelation function. The autocorrelation shows the extent to which events in one year are independent of those in neighbouring years out to a specified number of lagged years. Figure 4.5 shows the autocorrelation functions out to a lag of 10 years for the seasonal mean NA, EU and NH SCA indices 1972-2002. Generally, the ECHAM4 and observed autocorrelations correspond well, particularly at lag 1 year. There are greater differences during summer, with ECHAM4 underestimating the observed lag-1 autocorrelation by 0.5 over NA, 0.2 over EU and 0.25 for the NH. Furthermore, during summer the observed autocorrelation remains greater than 0.25 out to lag-5 for all regions. This suggests a high degree of multi-annual persistence in observed summer SCA that is underestimated by ECHAM4. By contrast, ECHAM4 does not capture the strong negative autocorrelation observed during winter over EU and NH. This feature in observations may warrant further investigation.

To assess fully the temporal SCA variability within ECHAM4, it is also instructive to examine the interseasonal persistence of SCA anomalies. Table 4.3 shows r_{RANK} between SCA anomalies over NA, EU and NH in summer (JJA) and winter (DJF) and between DJF and JJA 1972-2002. ECHAM4 underestimates persistence in both seasons over NA but overestimates persistence over EU. Larger errors occur over NA, where strong winter to summer persistence is seen in observations. Over the entire NH, ECHAM4 overestimates slightly the persistence in both seasons, which is due to the dominance of the EU index in NH SCA (Table 4.2). The relationship linking summer to winter climate is examined in detail in Chapter 5.

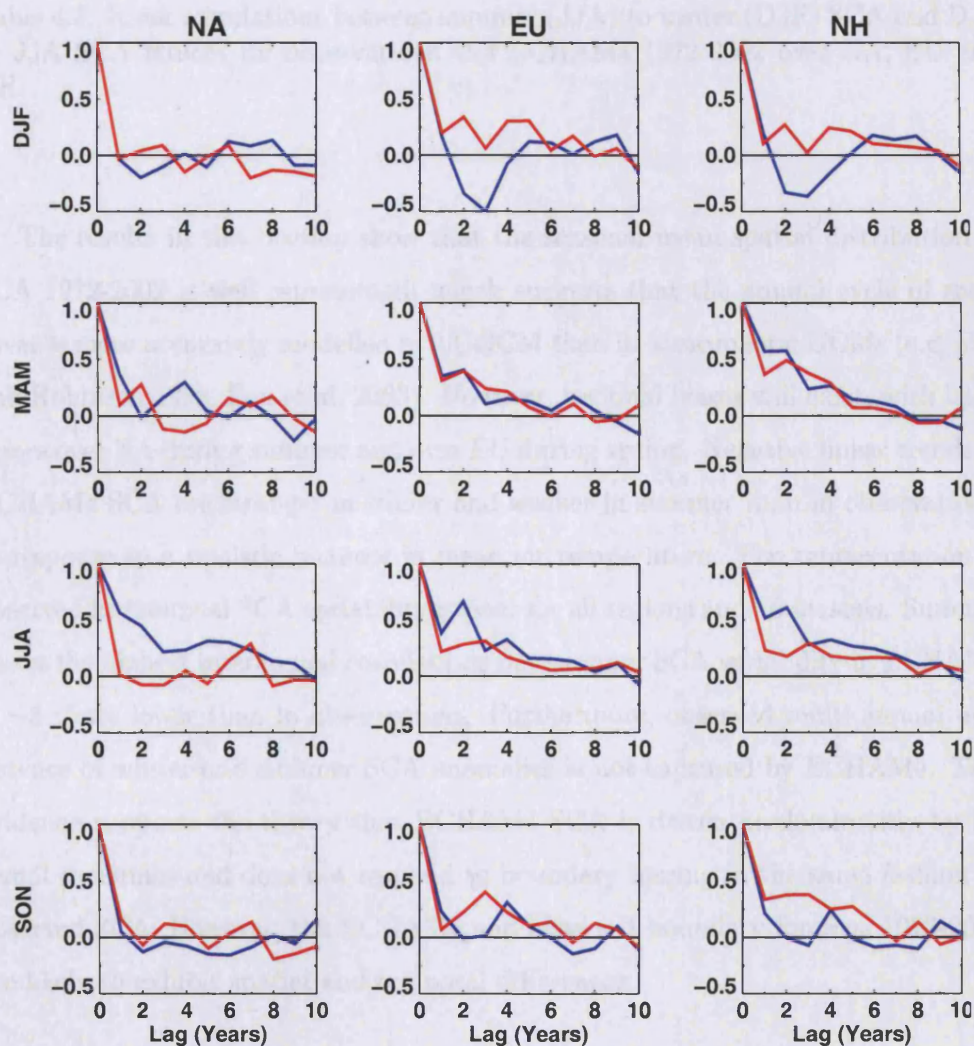


Figure 4.5: Autocorrelation function out to a time lag of 10 years for seasonal NA, EU and NH SCA indices 1972-2002. Blue line is observed SCA and red line is ECHAM4 SCA.

Region	OBS		ECHAM4	
	JJA-DJF	DJF-JJA	JJA-DJF	DJF-JJA
NA	0.16	0.39	-0.14	0.09
EU	-0.09	0.06	0.22	0.15
NH	0.11	0.19	0.26	0.20

Table 4.3: Rank correlations between summer (JJA) to winter (DJF) SCA and DJF to JJA SCA indices for observations and ECHAM4 1972-2002 over NA, EU and NH.

The results in this Section show that the seasonal mean spatial distribution of SCA 1972-2002 is well represented, which suggests that the annual cycle of snow cover is more accurately modelled in a CGCM than in atmospheric GCMs (e.g, Frei and Robinson 1998, Frei et al. 2003). However, regional biases still exist, with large errors over NA during summer and over EU during spring. Negative linear trends in ECHAM4 SCA are stronger in winter and weaker in summer than in observations, in response to a realistic increase in mean air temperature. The representation of observed interannual SCA variability is poor for all regions and all seasons. Summer shows the highest interannual correlations but summer SCA variability in ECHAM4 is ~ 3 times lower than in observations. Furthermore, observed multi-annual persistence of winter and summer SCA anomalies is not captured by ECHAM4. This evidence supports the theory that ECHAM4 SCA is driven predominantly by internal dynamics and does not respond to boundary forcing in the same fashion as observed SCA. However, the ECHAM4 and observed boundary forcings 1972-2002 are likely to exhibit spatial and temporal differences.

4.5 Climatic response to *in situ* snow anomalies

We now assess the regional climatic response associated with changes in contemporaneous *in situ* SCA over EU and NA. As discussed in Section 1.5.2, a fundamental problem exists in diagnosing cause and effect from contemporaneous snow cover analyses. Therefore, no assumptions can be made about the dominant direction of any association, i.e., from snow to atmosphere or from atmosphere to snow. How-

ever, we are concerned with assessing the differences between the observed and CGCM depictions of the snow/atmosphere relationship. The term ‘response’ is therefore used in the following only to describe an association between snow and the atmosphere and does not imply causality.

4.5.1 Atmospheric temperature response

Figure 4.6 shows the observed and ECHAM4 zonal mean atmospheric temperature response as a function of height over EU (50°E-100°E) associated with high minus low contemporaneous EU SCA terciles. The nature of the atmospheric response was found to be approximately linear (not shown), which validates the use of the composite difference between high and low snow events. Positive anomalies indicate a warming and negative anomalies indicate a cooling associated with high SCA anomalies. The signs are reversed for low SCA anomalies.

In observations, during all seasons a cooling is seen above the snow covered region that extends barotropically to the tropopause level. The cooling is strongest (and most significant) closer to the surface below 850 hPa, where magnitudes are $\sim 3^{\circ}\text{C}$. During winter and spring, when SCA is most abundant over the EU region, the cooling involves the entire land area north of 40°N . A less coherent atmospheric response is observed at $\sim 30^{\circ}\text{N}$. During summer, the cooling is confined to the region north of 60°N , although it is not confined in the vertical. The atmospheric pattern for autumn is more complicated, with a weaker (not significant) cooling centred on 55°N and a significant warming to the north at $\sim 75^{\circ}\text{N}$. During all seasons, the stratospheric response is opposite-signed to the tropospheric response, and of similar magnitude.

In ECHAM4, the overall magnitude and spatial pattern of the EU temperature response generally matches the observations winter through summer and shows greatest differences during autumn. The autumn pattern shows an overly significant cooling and no warming at $\sim 75^{\circ}\text{N}$. There are also certain differences between the patterns winter through summer. First, during winter the cooling in ECHAM4 is centred $\sim 10^{\circ}$ further south than observations even though the spatial extent of

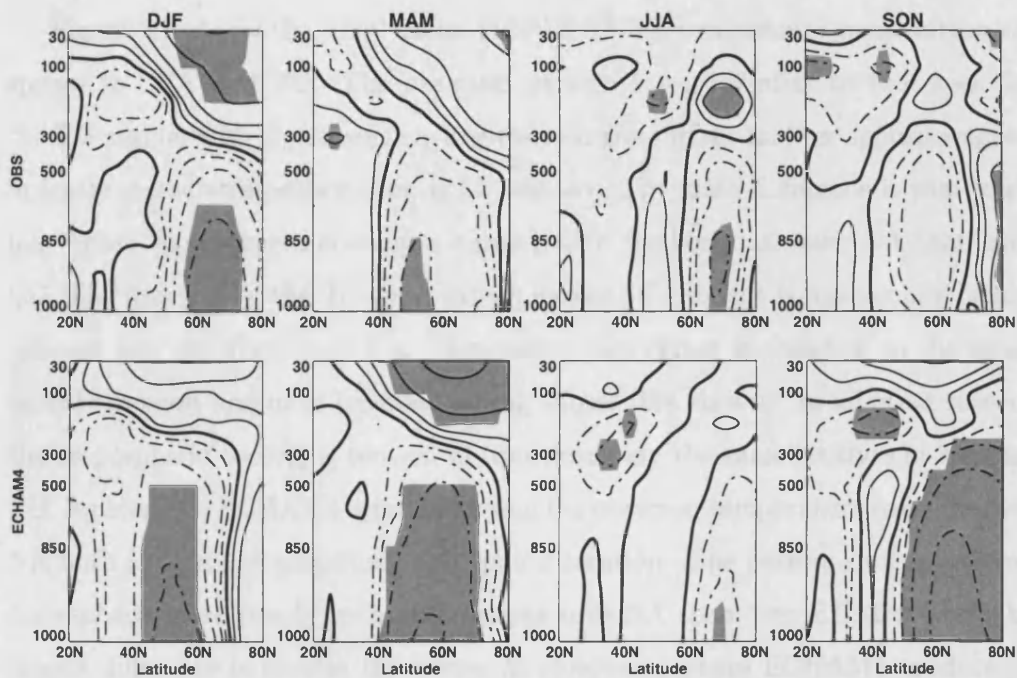


Figure 4.6: Zonally averaged (50°E - 100°E) Eurasian atmospheric temperature on pressure levels for high minus low EU SCA terciles for observations (NCEP/NCAR reanalysis, top panels) and ECHAM4 (bottom panels) at latitudes 20°N - 80°N . Data are temperature anomalies and contours are plotted at 0.25, 0.5, 1, 2 and 4°C . Dashed contours indicate negative anomalies and the thick solid contour denotes the zero line. Shading denotes significant ($p < 0.05$) differences as determined by a Student's t -test.

SCA is highly similar to observations (Figure 4.1). This suggests that ECHAM4 responds more strongly to SCA over Kazakhstan and southern Siberia compared to the observations, which appear to respond more to SCA over central Siberia. Second, during spring the significance of the tropospheric cooling is overestimated compared to observations. Third, during winter and summer the magnitude of the ECHAM4 stratospheric response is $\sim 25\%$ lower than in observations.

Figure 4.7 shows the zonal mean (120°W - 90°W) atmospheric temperature response to SCA over NA. The observed pattern is very similar to that seen for the EU region, with significant cooling close to the surface and an opposite-signed anomaly in the stratosphere during all seasons. The main difference is that, during winter, the observed cooling is centred $\sim 10^\circ$ further south over NA than over EU. The presence of the Tibetan plateau means SCA extent is less zonally homogeneous over EU than over NA. The cooling over Tibet is (relative to the zonal mean) strongest and most localised during winter (not shown). In all other seasons the tropospheric cooling is centred at approximately the same latitude in NA and EU. As over EU, ECHAM4 reproduces well the observed temperature response over NA both in terms of magnitude and spatial location. The pattern during autumn corresponds more closely with observations over NA than over EU. However, the largest difference is seen in the winter stratosphere, where ECHAM4 produces a significant high latitude cooling where the observations show a warming.

The atmospheric response to SCA during summer over both EU and NA is interesting for two reasons. First, during summer the SCA is confined to relatively small areas of high elevation over Tibet, northern Siberia and northern and western Canada (see Figure 4.1). Despite this limited extent, the observed temperature response remains significant during summer through the depth of the atmosphere. Second, despite low observed persistence between winter and summer (Table 4.3) there is strong persistence from spring to summer ($r_{\text{RANK}} = 0.79$). Mean summer snowfall over EU and NA is almost zero away from the highest peaks and therefore the majority of the SCA in summer is likely to have fallen during spring and persisted. This is an important finding because it suggests that the significant atmospheric signal observed during summer is the result of a lagged relationship with

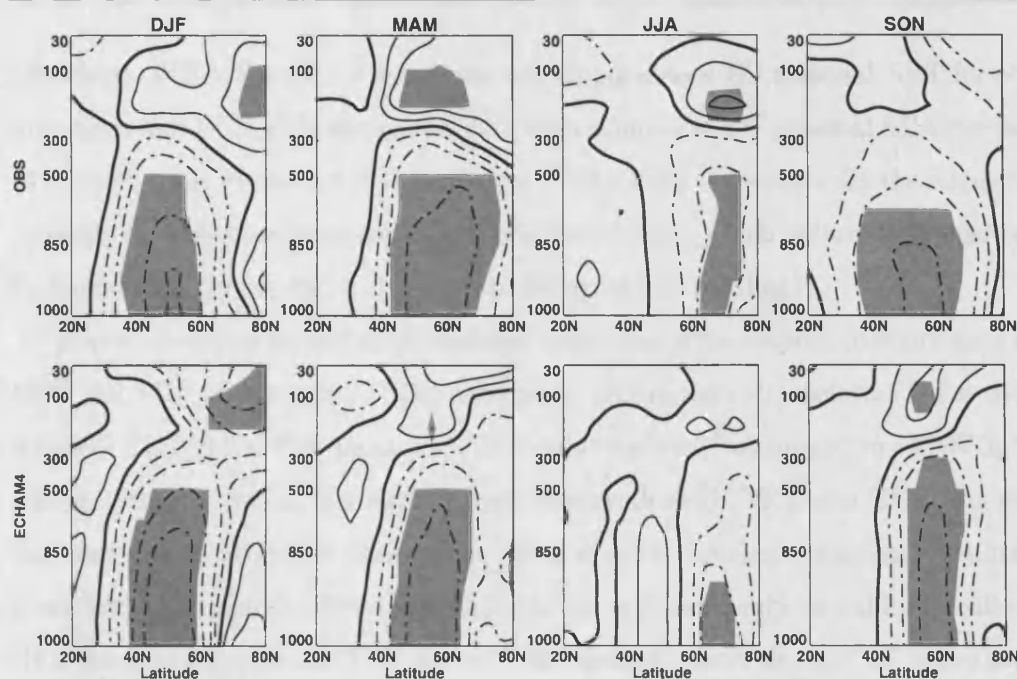


Figure 4.7: As Figure 4.6 except for zonally averaged (120°W - 90°W) NA atmospheric temperature, NA SCA terciles and contours are plotted at 0.5, 1, 2, 3 and 4°C .

spring snowfall. Furthermore, this persistence is captured accurately by ECHAM4.

In summary, the *in situ* atmospheric response to SCA is consistent between seasons and continents. High SCA terciles are associated with an equivalent barotropic cooling, which extends from the surface to the tropopause above the snow covered regions. An opposite-signed anomaly is seen in the stratosphere. ECHAM4 reproduces these observed patterns with accuracy in spatial position and magnitude, although differences are most pronounced during winter and autumn in both EU and NA.

4.5.2 Surface flux response

As discussed in Section 1.5.1, snow interacts with the overlying atmosphere predominantly through its effects on the surface energy balance. We now examine these effects through net radiative fluxes (net short-wave plus net long-wave radiation, henceforth NET) and net turbulent fluxes (sensible plus latent heat fluxes,

henceforth TUR). Figure 4.8 shows the contemporaneous EU seasonal NET for observations and ECHAM4 associated with high minus low EU seasonal SCA terciles 1972-2002, while Figure 4.9 shows seasonal TUR for the same years. In the standard convention, fluxes are positive away from the surface, which means that positive flux anomalies correspond to the surface losing energy (cooling).

The observations for EU show seasonal variations in the respective dominance of NET and TUR in response to SCA anomalies. During winter, insolation is low over northern EU (typical DJF mean at 60°N is only $\sim 50 \text{ Wm}^{-2}$ compared to $\sim 350 \text{ Wm}^{-2}$ during spring). Overall the surface loses energy through NET and TUR but the flux magnitudes are lower than in the other seasons. During spring and autumn, when both SCA and insolation are high, the snow covered regions exhibit localised NET gains and large-scale TUR losses. The summer season shows NET losses over much of EU and strong TUR losses over the permanent snow covered regions of northeast Siberia.

The NET and TUR fluxes in ECHAM4 for EU show little resemblance to those in observations for any season. The Pearson pattern correlations between observed and ECHAM4 NET and TUR are shown on the left of Table 4.4. None of the pattern correlation coefficients are significant and the strongest coefficient is negative ($r = -0.21$). During winter, ECHAM4 correctly simulates the low NET fluxes over northern EU but overestimates NET losses over southern EU. Over snow cover the majority of ECHAM4 fluxes occur through NET during all seasons, whereas in observations TUR is dominant.

Figure 4.10 shows the contemporaneous NA NET anomalies associated with high minus low NA seasonal SCA terciles, while Figure 4.11 shows the TUR fluxes for the same years. The observations show NET losses in all seasons except summer, when large NET gains are seen across NA. During winter, observations show simultaneous losses in NET and TUR over central NA, which suggests greater cooling than over EU for the same season. For the other seasons, NET dominates over NA, with TUR anomalies more localised and weaker in magnitude. In general, the observed flux response over NA is more coherent than over EU, which suggests the surface energy budget responds to SCA in a more uniform fashion. This could result from

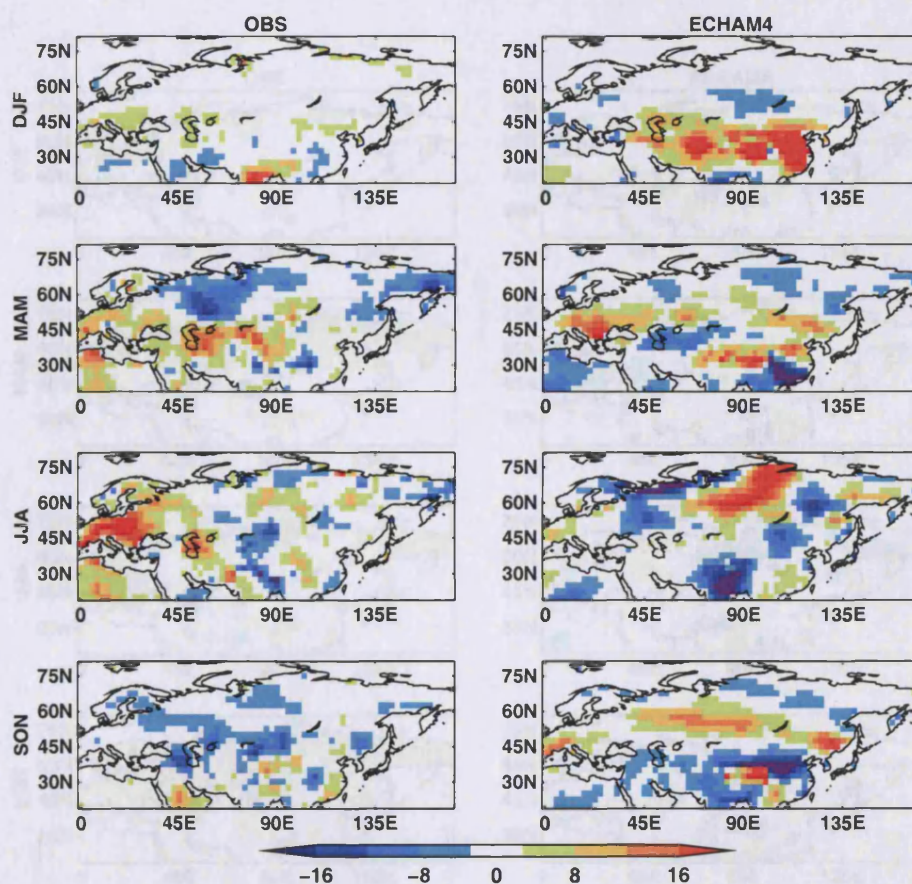


Figure 4.8: Observed (NCEP/NCAR reanalysis, left) and ECHAM4 (right) seasonal mean surface net radiative flux anomalies (Wm^{-2}) over EU for high minus low EU SCA terciles 1972-2002. Colours indicate positive (reds, surface losing energy) or negative (blues, surface gaining energy) fluxes.

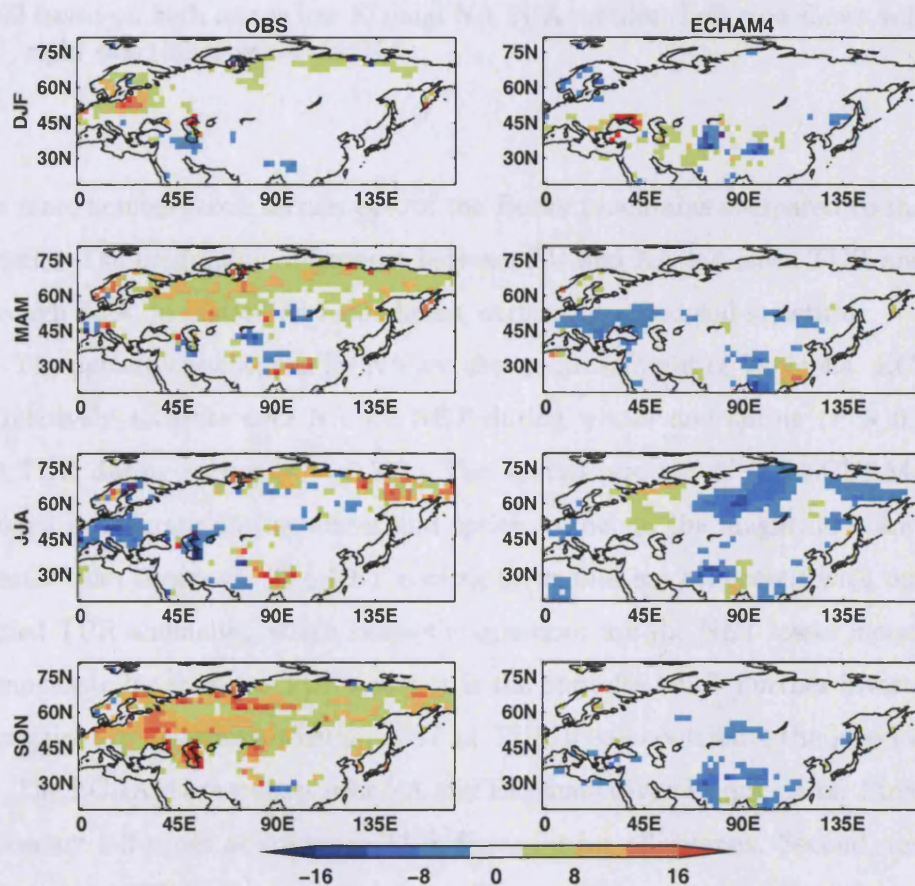


Figure 4.9: As Figure 4.8 except for sensible plus latent heat fluxes.

Season	Eurasia		North America	
	NET	TUR	NET	TUR
DJF	-0.12	-0.10	0.59	0.15
MAM	-0.02	0.07	0.57	0.37
JJA	0.06	-0.09	0.15	0.05
SON	-0.21	0.13	-0.08	0.24

Table 4.4: Pearson pattern correlations between seasonal mean net surface radiative (NET) and sensible plus latent (TUR) fluxes for observations and ECHAM4 1972-2002 based on high minus low EU and NA SCA terciles. Left side shows values for EU, right side shows values for NA.

the more homogeneous terrain east of the Rocky mountains compared to that over Siberia. The orographic differences between EU and NA influence TUR and NET through SCA, insolation and cloudiness, surface type and soil moisture.

The pattern correlations for NA are shown on the right of Table 4.4. ECHAM4 is relatively accurate over NA for NET during winter and spring ($r > 0.5$) and for TUR during spring ($r = 0.37$). The spatial position of the ECHAM4 NET cooling is accurate during winter and spring, although the magnitudes are $\sim 50\%$ greater than observed. The NET cooling anomalies are collocated with opposite-signed TUR anomalies, which raises the question: are the NET losses increased to compensate for spurious TUR gains or is the opposite true? Further investigation is required to determine whether NET or TUR fluxes contribute the larger errors.

The ECHAM4 responses over NA and EU share several similarities. First, NET fluxes are 2-3 times as strong as TUR fluxes during all seasons. Second, the magnitudes of NET fluxes are overestimated compared to observations, particularly during winter and spring. Third, TUR fluxes over snow cover in ECHAM4 are almost always negative, whereas in observations they are almost always positive. This difference is seen most clearly during summer, when the SCA is confined to northern Canada and Siberia (see Figure 4.1). In ECHAM4, summer TUR gains over the snow covered regions balance NET losses, whereas in observations the opposite occurs. This suggests that the largest errors in the ECHAM4 response to SCA occur *in situ* above the snow cover.

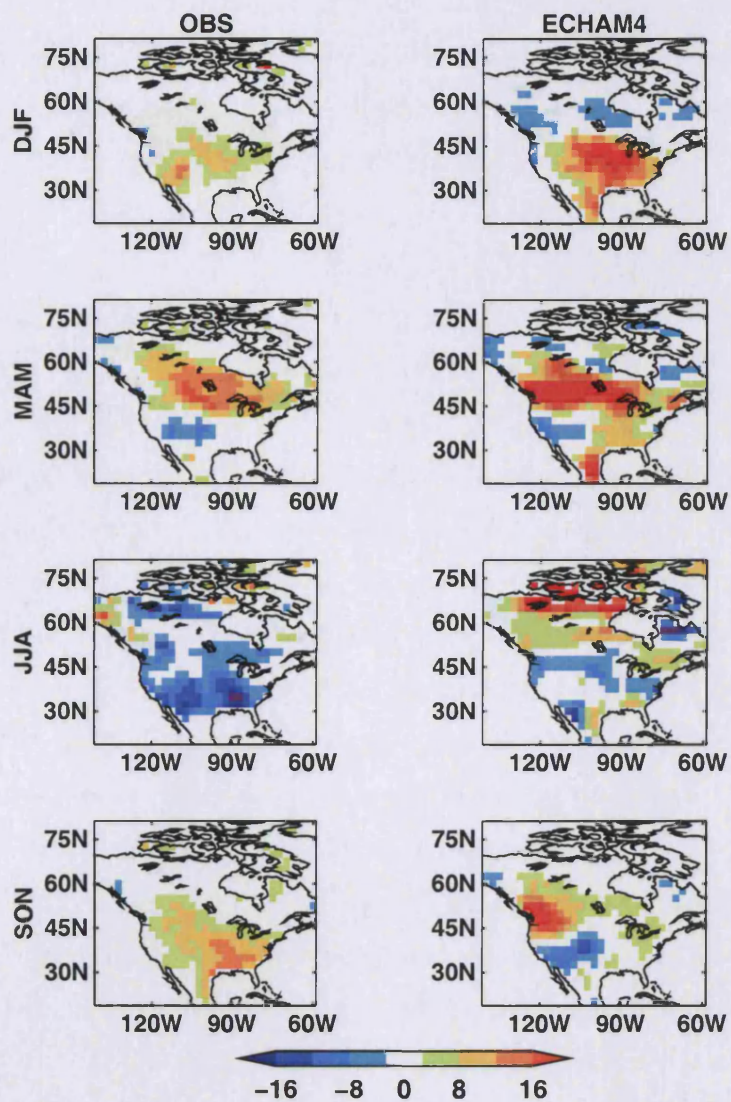


Figure 4.10: As Figure 4.8 except for NA net surface radiative fluxes and SCA anomalies.

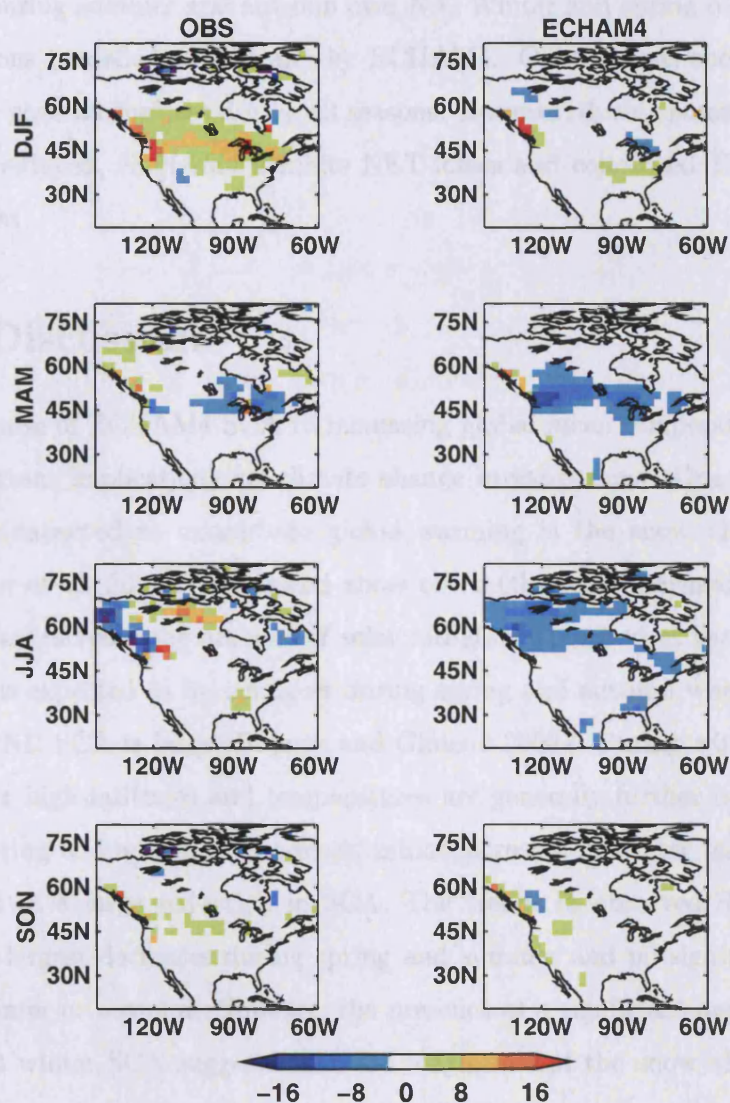


Figure 4.11: As Figure 4.8 except for NA sensible plus latent heat fluxes and SCA anomalies.

In summary, large differences are found between observations and ECHAM4 in both the magnitude and spatial pattern of the NET and TUR flux responses to anomalous *in situ* SCA. The differences are most pronounced for all seasons over EU and during summer and autumn over NA. Winter and spring over NA are the only seasons modelled realistically by ECHAM4. Observations show TUR losses over snow covered surfaces during all seasons. However, during summer when SCA extent is reduced, ECHAM4 exhibits NET losses and collocated TUR gains over snow cover.

4.6 Discussion

The response of ECHAM4 SCA to increasing global mean temperature 1972-2002 has important implications for climate change investigations. One of the climate feedbacks expected to exacerbate global warming is the snow albedo feedback (Houghton et al. 2001). Decreased snow cover (through increased melting) will lower α and increase the amount of solar radiation absorbed at the surface. This feedback is expected to be strongest during spring and autumn when insolation is high and NH SCA is large (Bojariu and Gimeno 2003). During winter, insolation is small at high latitudes and temperatures are generally further below 0°C than during spring and autumn. Therefore, minor increases in winter temperature will not result in a large reduction in SCA. The trends in observed SCA 1972-2002 show the largest decreases during spring and summer and no significant decrease during winter or autumn. However, the presence of a significant negative trend in ECHAM4 winter SCA suggests that the magnitude of the snow albedo feedback is overestimated compared to observations. In the shorter-term, this has implications for seasonal predictability because the resulting trends in ECHAM4 winter temperatures may act to damp or eradicate predictive links.

The NET and TUR flux responses to SCA in ECHAM4 show large differences to observations over EU and NA. There are several factors that could explain these errors. First, seasonal mean $\alpha \sim 0.7$ over snow covered land poleward of 60°N. This is realistic for fresh snow but somewhat high for snow that has been lying on

the ground for more than a few weeks. Second, atmospheric transmissivity (τ , the amount of short-wave radiation reaching the surface) may be inaccurate. Seasonal mean surface net short-wave radiation was found to be realistic (not shown), which means that τ could be overestimated to compensate for the high values of α . Third, clouds have a net warming effect at high latitudes because the decrease in insolation is outweighed by an increase in long-wave radiation from the clouds to the surface. Further investigation is required into the reciprocal influences of SCA and cloud cover in ECHAM4. This could examine the relationship between SCA and NET fluxes at the top of the atmosphere during clear and cloudy sky conditions (e.g., Groisman et al. 1994). However, cloud processes are heavily parameterised in ECHAM4 and additional complications could arise from the introduction of the indirect sulphate aerosol effect in GSDIO (Roeckner et al. 1999).

In observations and (to a lesser degree) in ECHAM4, NET cooling anomalies are located further south during autumn than during spring, when the SCA extents are similar (Figure 4.1). One explanation for this could be that soil at the snow margins is wetter during spring because of the moisture released by winter snow melt. The moisture would lower α in spring relative to the drier soils during autumn. Soil moisture is known to play a large role in snow feedback on the surface energy budget (Ose 1996). Further investigation is required to ascertain the effects of soil moisture from melting snow in ECHAM4. This could help to explain some of the inaccuracies in the surface flux response.

4.7 Summary and conclusions

This Chapter presents a comparison of the temporal and spatial variability of seasonal SCA 1972-2002 in observations and a coupled twentieth century climate simulation. SCA indices for the coupled GCM (ECHAM4), were computed as the fraction of the snow covered land area from monthly mean snow depth fields. Negative linear trends in ECHAM4 SCA are stronger than observations in winter and weaker in summer. These trends occur in response to a realistic increase in mean air temperature, which may cause overestimation of the winter snow/albedo feedback

in ECHAM4. The annual SCA cycle in ECHAM4 compares closely with observations. However, seasonally varying regional biases are found, which are largest over NA during summer and over EU during spring. The representation of observed interannual SCA variability is poor for all regions and all seasons. Compared to observations, ECHAM4 summer SCA shows highest interannual correlations but only one-third of the variability. Furthermore, observed multi-annual persistence of winter and summer SCA anomalies is not captured by ECHAM4. This evidence supports the theory that ECHAM4 SCA is driven primarily by internal dynamics and that twentieth century boundary forcing in ECHAM4 is different to observations.

The *in situ* atmospheric responses to SCA are compared and found to be consistent between seasons and continents. High seasonal SCA is associated with an equivalent barotropic cooling, which extends from the surface to the tropopause above the snow covered regions. An opposite-signed anomaly is seen in the stratosphere. ECHAM4 reproduces these observed patterns with accuracy in spatial position and magnitude, although differences are most pronounced during winter and autumn in both EU and NA. ECHAM4 captures accurately the significant lagged snow/temperature feedback in summer. However, ECHAM4 seasonal mean surface fluxes are accurate only over NA during winter and spring. The flux response above snow cover shows the largest errors in ECHAM4, particularly during summer.

Chapter 5

Coupled GCM Representation of Observed NAO_{DJF} Predictability from NH Summer Snow Cover

5.1 Introduction

In Chapter 4, the ECHAM4 CGCM run with observed historical radiative forcing was shown to reproduce accurately the observed annual cycle and spatial distribution of snow covered area (SCA) 1972-2002 over Eurasia (EU) and North America (NA). The seasonal *in situ* atmospheric temperature anomalies associated with SCA were also reproduced accurately. However, the interannual SCA variability in ECHAM4 compared poorly to observations. Evidence was presented suggesting that the ECHAM4 SCA indices were governed primarily by internal dynamics and responding to different boundary forcings. The surface radiative flux response over the snow covered regions was accurate only over NA during winter and spring. This showed that ECHAM4 did not capture realistically the modification to the surface energy budget caused by snow cover, which is the major influence that snow has on the atmosphere (Cohen and Rind 1991).

This Chapter investigates the ECHAM4 representation of observed links between Northern Hemisphere (NH) summer SCA and the winter NAO (NAO_{DJF}). We examine whether the CGCM captures the timing and magnitude of this relationship for the period 1972-2002. The teleconnected atmospheric and oceanic responses to changes in summer NH SCA are investigated. Saunders et al. (2003) proposed a feedback mechanism linking summer NH SCA and the NAO_{DJF}. This involved

a contemporaneous subpolar zonal temperature response to NH SCA, which is associated with circulation changes over the North Atlantic and anomalies in North Atlantic SST during subsequent months. Observational evidence supporting this mechanism is presented in Chapter 3 but the link appears to be nonstationary during the twentieth century. We use the ECHAM4 data available for the extended period 1904-2002 to assess the stationarity in this link.

This Chapter is structured as follows. The lead/lag correlation procedure is outlined in Section 5.2. Section 5.3 examines ECHAM4 twentieth century NAO_{DJF} variability. In Section 5.4 the NAO predictability arising from NH SCA within ECHAM4 is investigated and quantified for all leads and lags over the period 1972-2002. Section 5.5 presents an analysis of the physics underpinning the link between summer NH SCA and NAO_{DJF} in ECHAM4. This includes an investigation into the role of subpolar zonal temperature gradients in ECHAM4 NAO_{DJF} predictability and their teleconnection to the North Atlantic sector. The stationarity of the predictive snow/ NAO_{DJF} link and mechanisms is examined in Section 5.6. A discussion of the results is presented in Section 5.7 and the strengths and weaknesses in ECHAM4 for coupled snow/atmosphere investigations are highlighted in Section 5.8. Recommendations for potential improvements to CGCM formulations are made. Finally, Section 5.9 provides a brief summary of the main findings in this Chapter.

5.2 Methodology

5.2.1 Lead/lag correlation analysis

Spearman rank correlations are computed between SCA and the NAO for each month at every lead and lag between 0 and 12 months. In this way a matrix of correlation at all leads and lags is constructed. Each point represents the correlation between SCA data in a given month with the NAO index leading or lagging by up to 12-24 months. True physical correlations may be expected to extend to neighbouring cells in the correlation matrix. This is because physical relationships

would be expected to exert an influence for a time period greater than a single month. A subjective method of identifying true correlations in these matrices, as opposed to Type I errors, is therefore to examine clusters of locally significant ($p < 0.05$) correlations. An objective method is also employed, namely, the field significance test described in Section 3.2.3. In this Chapter, when correlation or regression is used to investigate predictability, a linear trend is removed from all data (see Section 3.2.2). This is consistent with the method of Saunders et al. (2003) and that employed in Chapter 3.

5.3 ECHAM4 NAO_{DJF} variability

Ahead of investigating the relationship between NH SCA and the NAO_{DJF} , it is instructive to assess NAO_{DJF} variability in ECHAM4. If the CGCM NAO_{DJF} bears no resemblance to the observed NAO_{DJF} , then even a perfectly modelled snow/atmosphere relationship may not offer predictability for NAO_{DJF} . Figure 5.1 shows the interannual variability in three ECHAM4 NAO_{DJF} indices 1904-2002. The ECHAM4 indices are calculated using the same methods and data regions employed in constructing the observed CRU, Hurrell and MSLP PC1 NAO_{DJF} indices used in Chapter 3 (see Section 3.2.1). Henceforth, ‘ECHAM4 CRU NAO_{DJF} ’ denotes the index computed using standardised MSLP from Gibraltar minus Iceland, while ‘ECHAM4 Hurrell NAO_{DJF} ’ employs standardised MSLP from Ponta Delgada (Azores) minus Iceland. ‘ECHAM4 MSLP PC1 NAO_{DJF} ’ denotes the index using the first principal component of North Atlantic sector MSLP. This means that the ECHAM4 NAO_{DJF} indices can be compared to the observational indices shown in Figure 3.1.

Several differences between the ECHAM4 and observed NAO_{DJF} indices are apparent. First, the correspondence between the three pairs of indices is low; the rank correlation between the observed CRU NAO_{DJF} and ECHAM4 CRU NAO_{DJF} is $r_{RANK} = 0.11$ for 1904-2002 and $r_{RANK} = 0.21$ for 1972-2002. Second, contrary to observations the linear trend is almost zero for all of the ECHAM4 indices 1904-2002. Over the shorter 1972-2002 period the trends are also small (not shown).

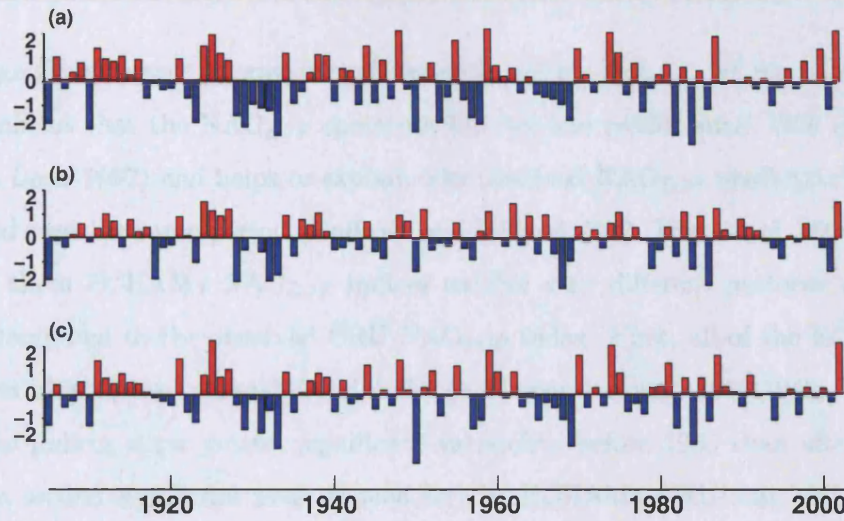


Figure 5.1: Temporal evolution of the ECHAM4 NAO indices. The indices are from ECHAM4 MSLP data and calculated using the same methods and data regions employed in constructing the observed indices shown in Figure 3.1, which are (a) CRU, (b) Hurrell and (c) MSLP PC1 NAO_{DJF} 1904-2002 in standardised units.

Third, the correspondence between the three ECHAM4 indices is lower than in observations. The highest correlation is found between the ECHAM4 MSLP PC1 and ECHAM4 CRU NAO_{DJF} indices ($r_{RANK} = 0.89$), while the lowest correlation is found between ECHAM4 Hurrell and ECHAM4 CRU NAO_{DJF} ($r_{RANK} = 0.69$). This compares with the range $r_{RANK} = 0.84$ to 0.85 in observations 1904-2002.

As discussed in Section 1.2.1, seasonal climate forecasting is concerned primarily with extracting low-frequency signals from climate data. Potential NAO_{DJF} predictability can be assessed through analysis of the dominant periods of variability in the NAO_{DJF} indices. If the NAO_{DJF} exhibits variability on time scales associated with boundary forcing then predictability may be expected. However, if the NAO_{DJF} exhibits variability consistent with an internal atmospheric process, then the likelihood of seasonal or interannual predictability is small. Figure 5.2 shows the wavelet power spectra and significance for the observed CRU and three ECHAM4 NAO_{DJF} indices 1904-2002. Areas of significant ($p < 0.05$) power are enclosed by a black contour line. The observed CRU index shows recent significant power at decadal periods: ~ 8 years 1970-1990 and ~ 6 years 1990-2000. Prior to 1970, there

is no significant power at any period except a minor peak at ~ 3 years in 1923. This confirms that the NAO_{DJF} spectrum has become redder since 1950 (Hurrell and van Loon 1997) and helps to explain why observed NAO_{DJF} predictability has increased over the same period (Rodwell and Folland 2002, Trigo et al. 2004).

The three ECHAM4 NAO_{DJF} indices exhibit very different patterns of variability compared to the observed CRU NAO_{DJF} index. First, all of the ECHAM4 indices exhibit a peak in decadal variability (~ 12 years) around 1910-1940. Second, all of the indices show greater significant variability before 1950 than after 1950. Third, a second significant peak is seen for the ECHAM4 CRU and MSLP PC1 NAO_{DJF} indices at a ~ 5 year period around 1950. Fourth, the spectrum for the ECHAM4 Hurrell index differs most from the other indices. The Hurrell index exhibits more power at shorter periods and the 1910-1940 decadal peak has greater significance. This shows a contrast to observations, where the spectra for the CRU, Hurrell and MSLP PC1 indices are in strong agreement (not shown).

These results highlight several reasons why ECHAM4 NAO_{DJF} may offer lower predictability than the observed NAO_{DJF} . First, the low correlation between the observed and ECHAM4 NAO_{DJF} indices means that they respond to twentieth century boundary forcing in different ways or that the boundary forcings are different. Observational NAO_{DJF} prediction schemes tend to perform better during periods of high decadal NAO variability (e.g., Rodwell and Folland 2002). Therefore, in observations greatest potential predictability occurs since 1970, while in ECHAM4 potential predictability is greatest during the early 1900s. Second, the lower correspondence between the three ECHAM4 NAO_{DJF} indices suggests that MSLP varies over smaller spatial scales in ECHAM4 than in observations. Third, the increasing greenhouse gas concentrations in GSDIO appear to have little impact on the NAO_{DJF} in ECHAM4. This concurs with the study of Paeth et al. (1999), who found small trends in annual mean NAO 1860-1989 but positive trends 1990-2060.

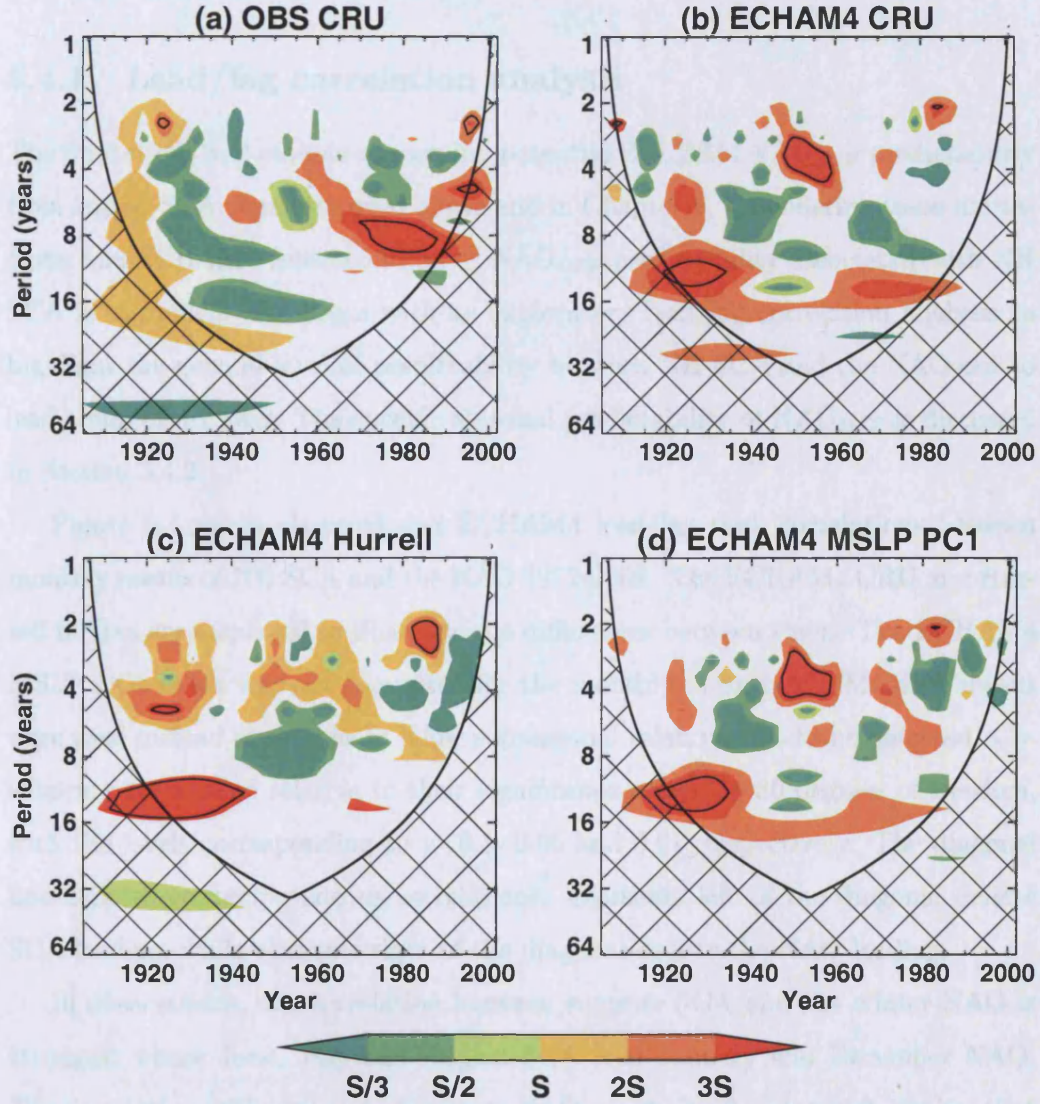


Figure 5.2: Wavelet power spectra of NAO_{DJF} indices 1904-2002. Plots show the observed (a) CRU, (b) ECHAM4 CRU, (c) ECHAM4 Hurrell and (d) ECHAM4 MSLP PC1 NAO_{DJF} indices. Cone of influence is indicated by cross-hatching. Colour shading indicates wavelet power at levels $S/3$, $S/2$, S , $2S$ and $3S$, where S is the standard deviation of the seasonal wavelet power. Black contour denotes significant ($p < 0.05$) power.

5.4 ECHAM4 NAO_{DJF} predictability from lagged NH snow cover

5.4.1 Lead/lag correlation analysis

The limitations and caveats concerning potential ECHAM4 NAO_{DJF} predictability from lagged SCA were discussed above and in Chapter 4. Considering these limitations, the ECHAM4 representation of NAO_{DJF} predictability associated with NH SCA is examined. We begin with an exploratory lead/lag correlation analysis to highlight the general level of predictability between NH SCA and the NAO out to lead times of ~ 1 year. The specific seasonal predictability of NAO_{DJF} is discussed in Section 5.4.2.

Figure 5.3 shows observed and ECHAM4 lead/lag rank correlations between monthly means of NH SCA and the NAO 1972-2002. The ECHAM4 CRU and Hurrell indices are displayed to illustrate the differences between them. The ECHAM4 MSLP PC1 index was not computed for the monthly mean NAO. Monthly means were used instead of seasons to allow subseasonal relationships to be detected. Coefficients are shaded relative to their significance assuming 30 degrees of freedom, with the levels corresponding to $p=0.1$, 0.05 and 0.01, respectively. The diagonal line denotes contemporaneous correlations. Elements left of the diagonal denote SCA leading, while elements right of the diagonal denote the NAO leading.

In observations, the correlation between summer SCA and the winter NAO is strongest where June, July and August SCA lead January and December NAO. The correlation with upcoming February NAO is not significant, which implies that summer SCA may be a better predictor of December-January mean NAO than NAO_{DJF}. There is also a significant correlation when January NAO leads June and July SCA, which indicates a link from summer to winter and then back to summer. Significant contemporaneous SCA/NAO correlations are confined to the winter months, when the NAO influence over EU SCA is at its peak (Robock et al. 2003). The NAO also leads significantly by one month SCA variations in March, May and October.

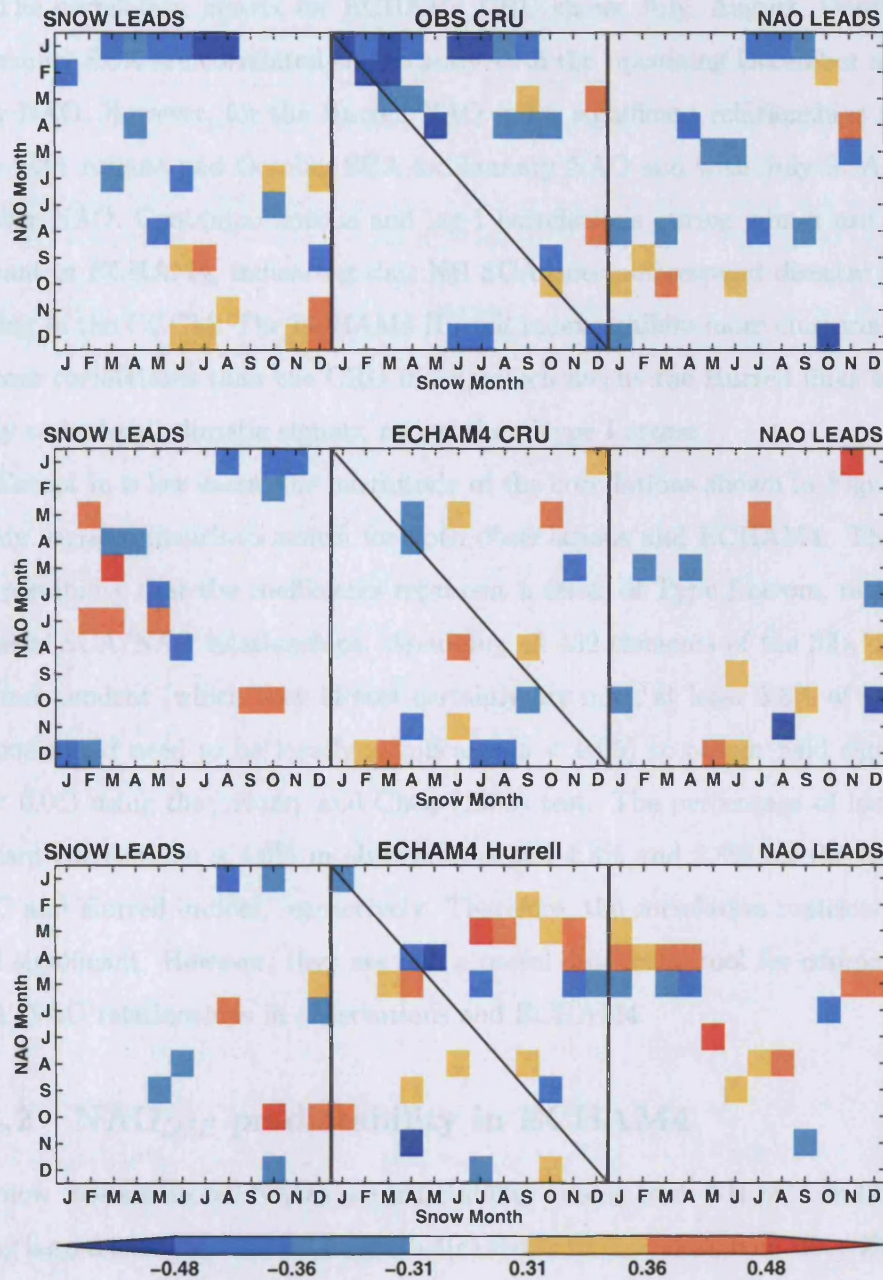


Figure 5.3: Lead/lag rank correlation matrices between monthly mean NH SCA and the NAO for observations (top), ECHAM4 CRU (centre) and ECHAM4 Hurrell (bottom). Contemporaneous correlations denoted by diagonal line. Elements left of diagonal denote SCA leading and right of diagonal denote NAO leading. Colour shading denotes statistical significance assuming 30 degrees of freedom at 0.01, 0.05 and 0.1 levels; blues indicate negative, while reds indicate positive correlation. Coefficients where $p > 0.1$ are not plotted. All data are detrended.

The correlation matrix for ECHAM4 CRU shows July, August, October and November SCA are correlated significantly with the upcoming December and January NAO. However, for the Hurrell NAO index significant relationships are seen only with August and October SCA for January NAO and with July SCA for December NAO. Contemporaneous and lag-1 correlations during winter are not significant in ECHAM4, indicating that NH SCA does not respond directly to NAO forcing in the CGCM. The ECHAM4 Hurrell index exhibits more clustering of significant correlations than the CRU index, which means the Hurrell links are more likely to be ‘real’ climatic signals, rather than Type I errors.

Except in a few cases, the magnitude of the correlations shown in Figure 5.3 is highly variable month-to-month for both observations and ECHAM4. This raises the possibility that the coefficients represent a series of Type I errors, rather than physical SCA/NAO relationships. Assuming all 432 elements of the 32×12 matrix are independent (which they almost certainly are not), at least 6.8% of the correlations would need to be locally significant ($p < 0.05$) to obtain field significance ($p < 0.05$) using the Livezey and Chen (1983) test. The percentage of locally significant correlations is 4.6% in observations and 4.3% and 3.8% for the ECHAM4 CRU and Hurrell indices, respectively. Therefore, the correlation matrices are not field significant. However, they are still a useful diagnostic tool for comparing the SCA/NAO relationships in observations and ECHAM4.

5.4.2 NAO_{DJF} predictability in ECHAM4

We now assess seasonal NAO_{DJF} predictability arising from NH SCA in ECHAM4 using lagged rank correlation. Figure 5.4(c) shows the correlation between ECHAM4 bi-monthly NH SCA and the three NAO_{DJF} indices 1972-2002 (panels (a) and (b) are discussed in Section 5.6.1). The significant link peaks in summer for July-August (JA) mean NH SCA, which is one month later than in observations (see Figure 1.6). A similar pattern was found using EU SCA but for NA SCA the summer peak was positive and not significant (not shown). The differences between the ECHAM4 NAO_{DJF} indices mean the CRU and Hurrell indices are borderline

Period	Lagged Predictor	CRU NAO _{DJF}				Hurrell NAO _{DJF}				MSLP PC1 _{DJF}			
		r	(p)	$MSSS$	(p)	r	(p)	$MSSS$	(p)	r	(p)	$MSSS$	(p)
1937-69	JA NH SCA	0.19	(0.11)	3	(0.16)	0.25	(0.06)	5	(0.09)	0.16	(0.12)	2	(0.19)
1972-2002	JA NH SCA	0.25	(0.07)	3	(0.09)	0	(-)	0	(-)	0.23	(0.10)	0	(-)

Table 5.1: Skill values from ECHAM4 cross-validated NAO_{DJF} hindcasts for 1937-69 (top) and 1972-2002 (bottom) using JA mean NH SCA. r is the correlation skill value and $MSSS$ is the percentage improvement in mean-square skill score over climatology. p is the probability that the observed skill value was obtained by random chance, with values shown in brackets. Skill significance levels determined by a Monte Carlo resampling test with 25,000 iterations. All data are detrended.

statistically significant. In general, MSLP PC1 slightly overestimates the positive correlation during spring, while Hurrell NAO_{DJF} underestimates the negative correlation during autumn. In contrast to observations (Figure 1.6), there is no significant contemporaneous link during winter between NH SCA and NAO_{DJF}. This shows that winter snow cover variations over the NH are not driven by the NAO_{DJF} circulation pattern. Instead, more regional-scale processes could be responsible for snowfall and persistence in SCA.

The significant link identified above between JA NH SCA and NAO_{DJF} 1972-2002 is quantified by constructing an empirical model and producing NAO_{DJF} hindcasts. The hindcast, skill verification and significance testing methodologies are the same as those employed in Chapter 3 (see Section 3.2.3). The cross-validated hindcast skill from ECHAM4 JA NH SCA for the three NAO_{DJF} indices 1972-2002 is shown in the bottom row of Table 5.1. The JA snow cover predictor produces positive, although not significant, skill against the CRU and MSLP PC1 NAO_{DJF} indices. However, there is zero skill for predicting the Hurrell index. This is an interesting finding given that the correlation with JA NH SCA is similar for all three NAO_{DJF} indices (Figure 5.4). This result, coupled with the relatively low skill for the CRU and MSLP PC1 indices, suggests that the physical link between JA NH SCA and NAO_{DJF} may be weaker in ECHAM4.

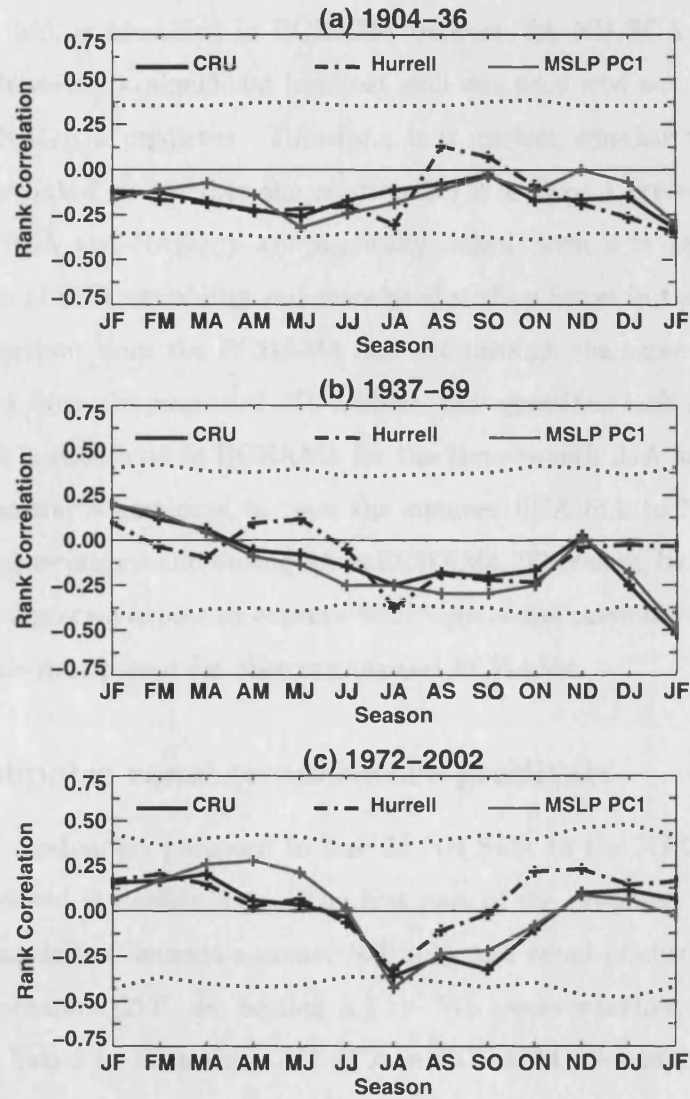


Figure 5.4: Lagged rank correlation between ECHAM4 bi-monthly mean NH SCA and the three NAO_{DJF} indices for (a) 1904-36, (b) 1937-69 and (c) 1972-2002. Dotted lines denote statistical significance ($p < 0.05$) corrected for serial correlation. All data are detrended.

5.5 Physical basis for NAO_{DJF} predictability in ECHAM4

A significant link is identified in ECHAM4 between JA NH SCA and NAO_{DJF} 1972-2002. However, no significant hindcast skill was produced using JA NH SCA as a lagged NAO_{DJF} predictor. Therefore, it is unclear whether these variables are physically linked or whether the relationship is a Type I error. If ECHAM4 summer NH SCA and NAO_{DJF} are physically linked, then it is despite the poor representation of SCA variability and associated surface fluxes in the CGCM. This raises the question: does the ECHAM4 link act through the same mechanism as that proposed from observations? To address this question, each element of the proposed link is examined in ECHAM4 for the three-month JJA summer season. This longer season is employed because the summer SCA link to NAO_{DJF} peaks during JJ in observations and during JA in ECHAM4. Therefore, by averaging over these three months we expect to capture both signals and maintain consistency by using the same time period for observations and ECHAM4.

5.5.1 Subpolar zonal temperature gradients

The physical mechanism proposed to link JJ NH SCA to the NAO_{DJF} in observations is outlined in Section 3.4.4. The first part of the proposed link involves a significant association between summer NH SCA and zonal gradients in subpolar 2 m air temperature (ΔT , see Section 3.2.1). We assess whether, as in observations, ΔT is linked to anomalous NH SCA in ECHAM4. Saunders et al. (2003) constructed ΔT based on three regions of significant temperature correlation with contemporaneous JJ NH SCA (Figure 3.2). Figure 5.5 shows the contemporaneous single-point rank correlation between JJA mean NH 2 m air temperature and NH SCA 1972-2002. Coefficients are shaded where $|r_{RANK}| > 0.2$ and are approximately significant ($p < 0.05$) where $|r_{RANK}| \geq 0.4$. For observations, the pattern is similar to Figure 3.2 except that correlation coefficients are plotted instead of correlation significance. The negative correlations over the snow covered regions of northern

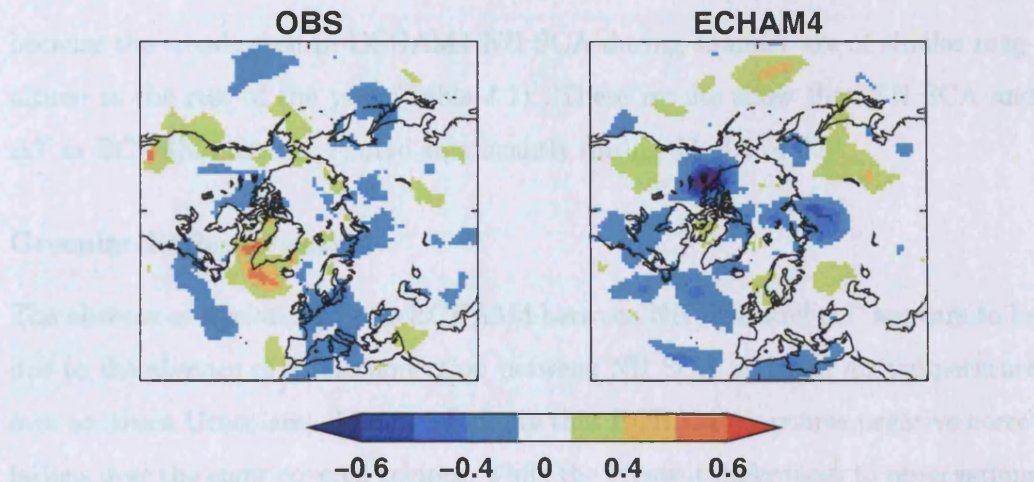


Figure 5.5: Rank correlation coefficients between JJA mean NH 2 m air temperature and NH SCA for observations (NCEP/NCAR reanalysis, left) and ECHAM4 (right) 1972-2002.

EU and NA are evident, as is the significant positive correlation over southern Greenland. The spatial pattern in ECHAM4 shows negative correlations overlying the snow in EU and NA, although correlation magnitudes are $\sim 50\%$ greater than observed. However, the teleconnected positive correlation over southern Greenland is absent in ECHAM4. Instead, the entire North Atlantic sector exhibits a negative relationship with NH SCA. This means that the subpolar zonal gradient in air temperature associated with NH SCA is not established in ECHAM4 during JJA.

A correlation analysis reveals that the absence of the zonal temperature gradients in ECHAM4 means that the JJA ΔT index is not correlated significantly with contemporaneous NH SCA (top half of Table 5.2). However, this could simply result from employing the JJA averaging period instead of JJ or JA. To test this, the contemporaneous relationships between NH SCA and ΔT for JJ and JA are also presented. In observations the link peaks in JJ and is $\sim 30\%$ stronger with detrended data than with raw data. This reflects the significant negative trends seen in summer SCA 1972-2002 (Table 4.1). ECHAM4 shows a similar pattern to observations but the coefficients are uniformly weaker and not significant during any season. Furthermore, trend has less impact than in observations, and detrending the ECHAM4 data only strengthens the correlation during JA and JJA. This is

because the trends seen in ECHAM4 NH SCA during summer are of similar magnitude to the rest of the year (Table 4.1). These results show that NH SCA and ΔT in ECHAM4 are not related significantly during JJ, JA or JJA.

Greenland teleconnection

The absence of a relationship in ECHAM4 between NH SCA and ΔT appears to be due to the absence of a teleconnection between NH SCA and 2 m air temperature over southern Greenland. Figure 5.5 shows that ECHAM4 captures negative correlations over the snow covered regions, while the greatest differences to observations occur over the North Atlantic sector. The Greenland region is weighted double compared to EU and NA in the ΔT index, and errors in this region also account for inconsistencies in observational ΔT analyses (see Section 3.3.5). If the absence of the Greenland teleconnection causes the NH SCA/ ΔT link to break down, then the link should be stronger using a ΔT index which excludes Greenland

$$\Delta_x T_{EG} = \frac{NA + EU}{2}, \quad (5.1)$$

where $\Delta_x T_{EG}$ represents ΔT excluding Greenland and EU and NA are the 2 m temperatures area averaged over the EU and NA regions, respectively, as shown in Figure 3.2. The correlations between NH SCA and ΔT_{EG} are listed in the bottom half of Table 5.2. The coefficients are significant in ECHAM4 with JJ and JJA ΔT_{EG} ; the latter is also significant using detrended data. Interestingly, the observed correlations are weaker with ΔT_{EG} than with ΔT . This reflects the importance in ΔT of the positive teleconnection to southern Greenland in Figure 5.5. These results show that the teleconnection between NH SCA and Greenland is modelled poorly in ECHAM4 and its absence has a significant impact on the relationship between NH SCA and ΔT .

Hemispheric response to NH snow cover

Next, we attempt to determine why the hemispheric-scale ΔT link does not become established in ECHAM4 during summer. This is achieved through a comparison

	OBS			ECHAM4		
	JJ	JA	JJA	JJ	JA	JJA
ΔT RAW	-0.41	-0.20	-0.35	-0.32	-0.10	-0.25
ΔT DET	-0.68	-0.41	-0.56	-0.32	-0.17	-0.32
ΔT_{EG} RAW	-0.28	-0.22	-0.26	-0.47	-0.30	-0.49
ΔT_{EG} DET	-0.43	-0.24	-0.33	-0.32	-0.16	-0.43

Table 5.2: Rank correlation coefficients between contemporaneous NH SCA and ΔT and ΔT_{EG} 1972-2002 for JJ, JA and JJA seasonal means. Bold type denotes significant ($p < 0.05$) correlations (corrected for serial correlation). RAW denotes data with trends included, while DET denotes data with a linear trend removed.

of the contemporaneous NH atmospheric circulation associated with NH SCA in observations and ECHAM4. Figure 5.6 shows JJA mean 500 hPa geopotential height anomalies poleward of 30°N associated with high minus low JJA NH SCA terciles 1972-2002. Observations show that the teleconnected response to NH SCA is larger in magnitude (and more significant) than the *in situ* response above the snow covered regions of EU and NA. South of 50°N there are continental-scale negative height anomalies across EU and NA. Away from the North Pole, large-scale positive height anomalies are found only over the North Atlantic south of Greenland. This corresponds to the region where warm air temperature anomalies are seen in Figure 5.5. The positive height anomalies adjacent to the North Pole are located upstream (west) of the EU and NA snow covered areas.

The ECHAM4 height response exhibits several differences to observations. First, a larger number of stronger, more localised height anomaly centres produce a less zonally homogeneous response than in observations. The ECHAM4 centres form a zonal ‘wave-train’ with positive (anticyclonic) height anomalies located adjacent to negative (cyclonic) height anomalies in series around the subpolar NH. However, this wave-train does not form a perfect circle and a weak meridional ridge extends from Scandinavia to the Middle East. Second, the response is stronger in ECHAM4 above the snow covered regions, particularly over northern Siberia. Third, the North Atlantic response south of Greenland in ECHAM4 has the opposite sign to observations, with the positive height anomaly located south of the negative

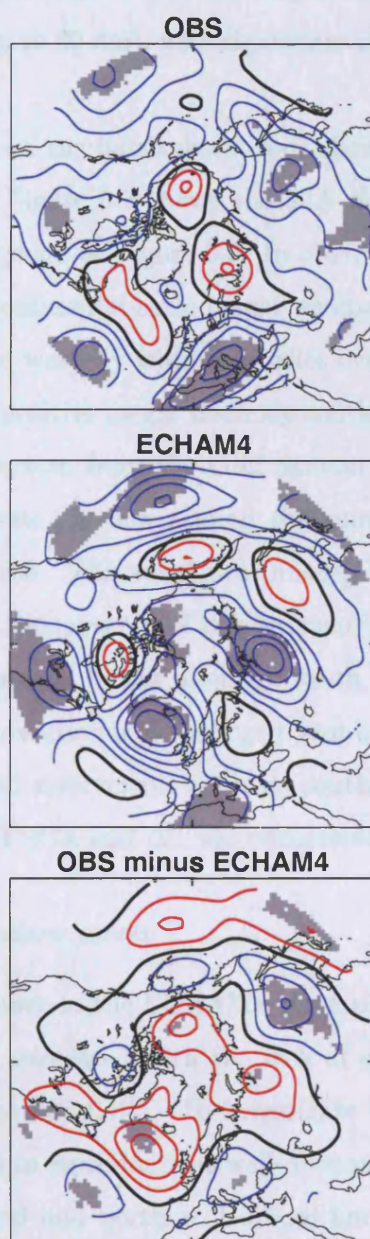


Figure 5.6: JJA mean 500 hPa geopotential height anomalies for observations (top), ECHAM4 (centre) and observations minus ECHAM4 (bottom) associated with high minus low JJA NH SCA terciles 1972-2002. Contour interval is 5 dm for top and centre panels and 10 dm for bottom panel. Red contours denote positive values and blue contours denote negative values. Grey shaded areas denote differences significant at 5% as determined by a Student's *t*-test.

anomaly. The difference plot shows that the largest height differences are found over the North Atlantic (up to 30 dm), with significant differences in the area south of Greenland.

The effect of NH SCA on the hemispheric zonal flow is best seen in the lower-tropospheric wind fields. Figure 5.7 shows the JJA mean 850 hPa wind vectors for the same composite of years as Figure 5.6. In observations the circulation pattern associated with the continental-scale height anomalies exhibits easterly wind anomalies over Eurasia and westerly wind anomalies over North America. The anticyclonic flow around the positive height anomaly south of Greenland is associated with warm, poleward advection from the Gulf Stream region. In ECHAM4, the 850 hPa wind pattern shows the flow around the smaller, more localised height anomalies seen in Figure 5.6. The zonal flow linking EU, NA and Greenland is not established because of the presence of ridges situated over Eastern Europe and northern Canada. The negative height anomaly south of Greenland is associated with cooler, equatorward advection into the region east of Newfoundland. Thus, the absence of the teleconnected anticyclonic anomaly south of Greenland in ECHAM4 is likely to explain why NH SCA and ΔT are uncorrelated.

Absence of Greenland snow cover

The problems identified above in the ECHAM4 atmospheric response near Greenland to NH SCA may be associated with the lack of snow cover over Greenland in the GSDIO run (see Section 4.4.1). To investigate this possibility we analyse the capability of ECHAM4 to reproduce the well-documented temperature teleconnection involving Greenland and northern Europe, known as the North Atlantic ‘seesaw’ (Rogers and van Loon 1979). This feature is strongest during winter but is observed all year round, albeit with weaker magnitude in summer. If ECHAM4 reproduces closely the observed seesaw pattern in the absence of Greenland snow, then it is reasonable to conclude that the lack of Greenland snow does not influence the atmospheric response to NH SCA.

The seesaw is represented by an index of monthly mean 2 m temperature area-

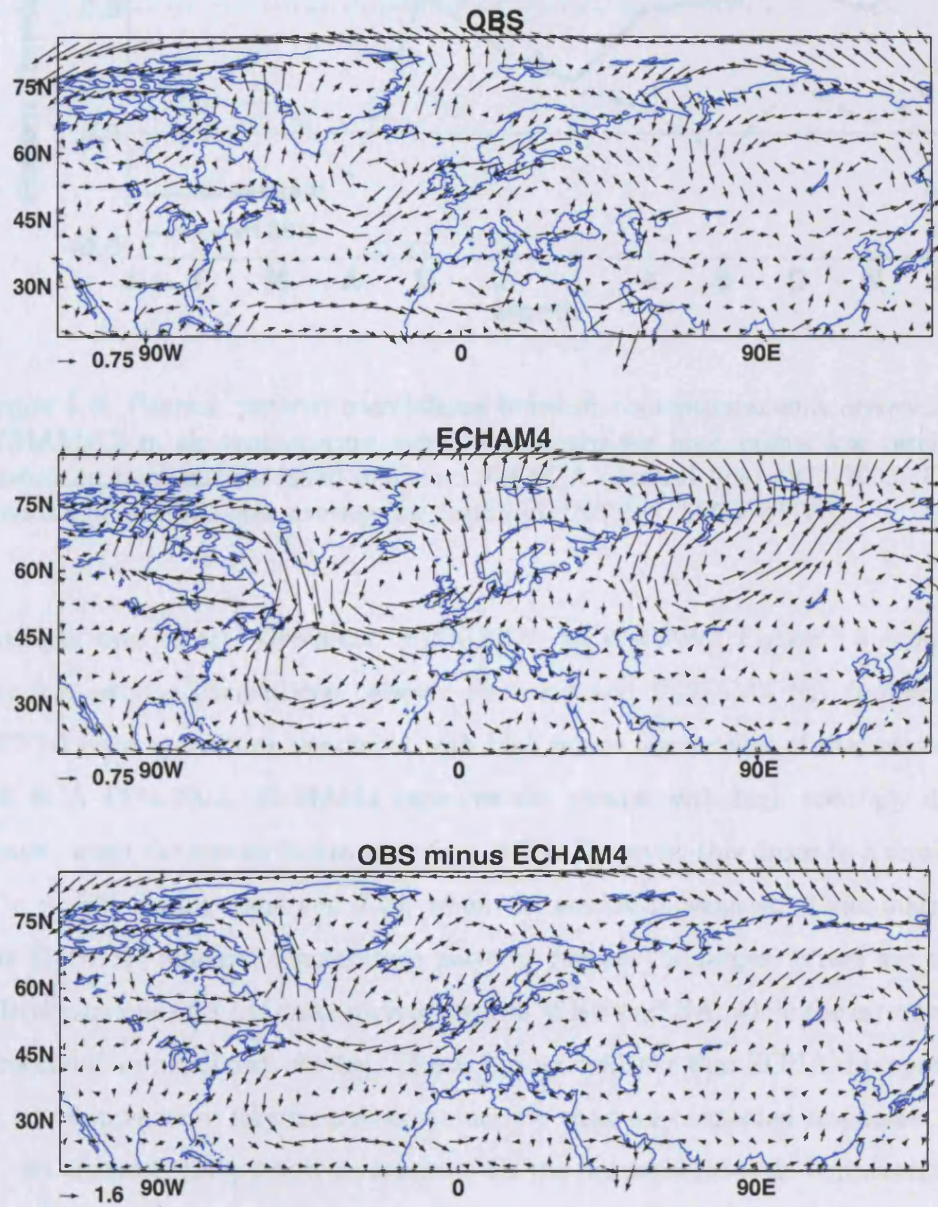


Figure 5.7: As Figure 5.6 except for JJA 850 hPa winds in observations (top), ECHAM4 (centre) and observations minus ECHAM4 (bottom). Vectors show monthly mean wind anomalies in ms^{-1} .

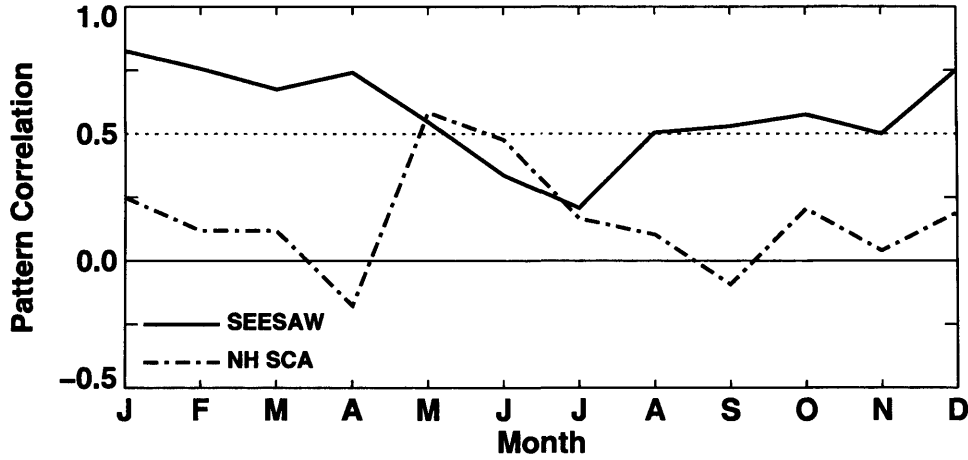


Figure 5.8: Pearson pattern correlations between contemporaneous observed and ECHAM4 2 m air temperature composite means for high minus low terciles of Greenland temperature (solid line) and NH SCA (dashed line) 1972-2002. Greenland temperature is area average for region [67°N-72°N, 52°W-47°W].

averaged over central Greenland [67°N-72°N, 52°W-47°W]. Figure 5.8 shows the monthly pattern correlations between observed and ECHAM4 NH (poleward of 50°N) 2 m temperatures associated with high minus low terciles of the seesaw and NH SCA 1972-2002. ECHAM4 captures the seesaw with high accuracy during winter, when the seesaw is strongest ($r \sim 0.75$). However, this drops to a minimum of $r \sim 0.25$ during June and July, when the seesaw is weakest. Examination of the ECHAM4 summer temperature patterns reveals the largest errors are in the teleconnections over the snow covered regions of EU and NA, while the errors above Greenland are small (not shown). This is further evidence that ECHAM4 represents *in situ* temperature responses more accurately than teleconnected responses.

By contrast, the pattern correlations for the hemispheric-scale temperature response to NH SCA are weaker than for the seesaw response in all months except May and June. Figures 4.6 and 4.7 showed that ECHAM4 depicts accurately the *in situ* temperature responses to SCA. However, on the hemispheric-scale the errors are up to $\sim 10^\circ\text{C}$; ECHAM4 underestimates the response to NH SCA for autumn through spring and overestimates the response during summer. During winter ECHAM4 is accurate in the temperature response above Greenland. However, ECHAM4 rep-

resents poorly the teleconnected responses to NH snow cover during all months. This suggests that errors in the ECHAM4 temperature response to NH SCA do not result from the absence of snow cover over Greenland.

5.5.2 North Atlantic sector response

The second part of the proposed mechanism linking NH SCA and NAO_{DJF} involves a contemporaneous atmospheric response to ΔT during summer centred on the North Atlantic. This leads by ~ 1 month patterns of SST near Newfoundland and in the subtropics, which persist into autumn. Since ΔT is not established during summer in ECHAM4, we examine instead the response to NH SCA over the North Atlantic sector. If the response in ECHAM4 without ΔT is the same as in observations with ΔT , this would suggest that ΔT does not play a major role in observed NAO_{DJF} predictability. Therefore, the absence of a snow/ ΔT link in ECHAM4 provides the opportunity to validate the observed ΔT link.

Figure 5.9 shows the evolution of ECHAM4 monthly mean anomalies in 850 hPa zonal wind, MSLP and SST over the North Atlantic sector April through October. The anomalies are associated with high minus low JJA NH SCA terciles 1972-2002. The SCA data are multiplied by -1 because the NH SCA and ΔT indices are negatively correlated (Table 5.2). This facilitates a comparison with Figure 3.6, which shows the response to ΔT terciles in observations 1950-2001. The zonal wind and MSLP responses show little coherency between months, which is reflected in the low magnitude JJASO seasonal means. Only during June and August is there any significant zonal wind response in ECHAM4 to NH SCA over the North Atlantic. In observations, there is a significant zonal wind response during both June and July and also in the JJASO mean.

The ECHAM4 SST fields show that significant SST anomalies evolve June through October in the absence of either ΔT or a consistent zonal wind or MSLP response. However, the SST anomalies exhibit several differences compared to those associated with ΔT in observations. First, the positive anomaly centred on $50^\circ N$ July through September is located northwest of the Newfoundland anomaly seen in

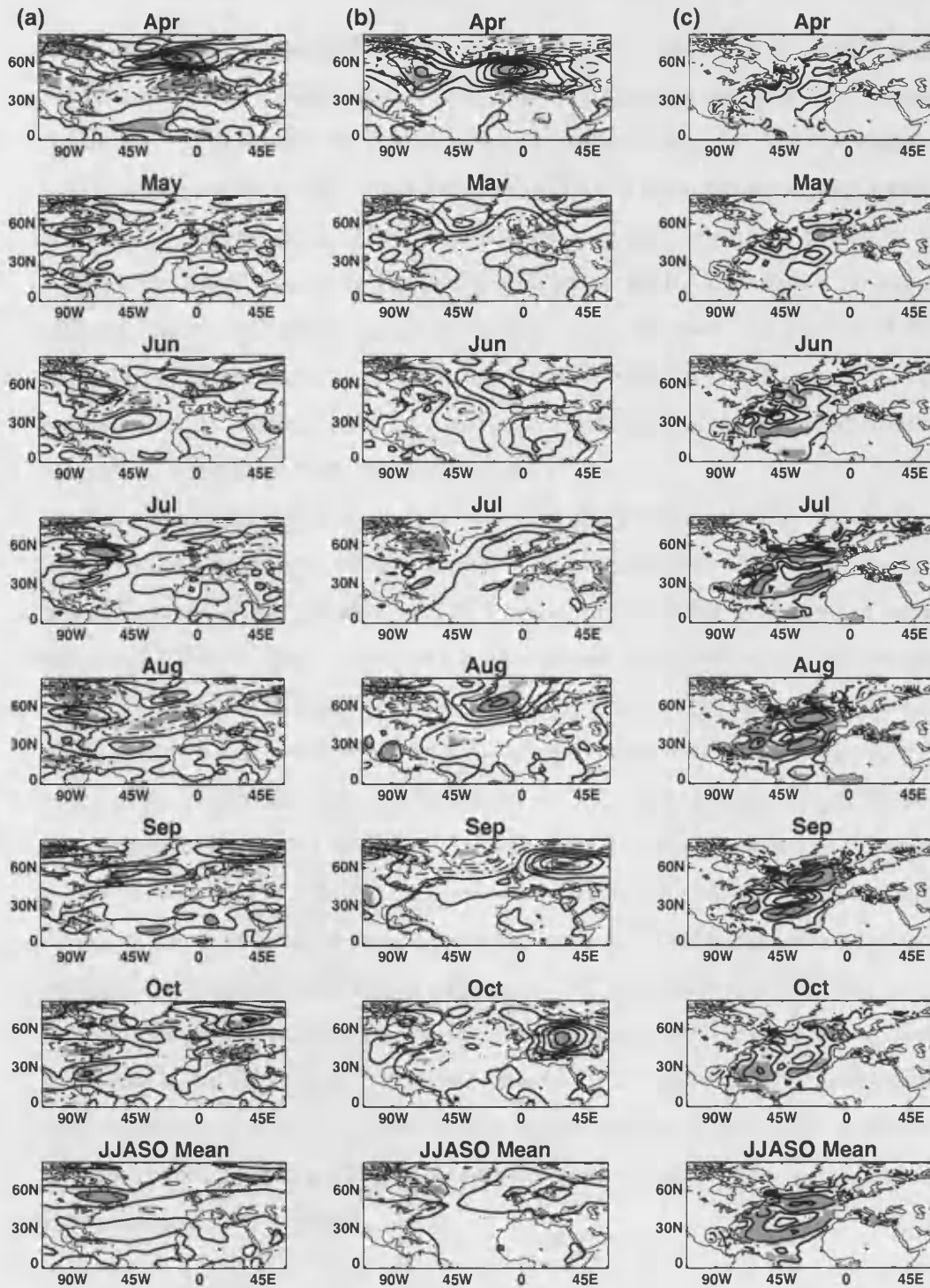


Figure 5.9: As Figure 3.6 except for ECHAM4 JJA mean (a) 850 hPa zonal wind, (b) MSLP and (c) SST anomalies associated with high minus low JJA NH SCA terciles 1972-2002. NH SCA index is multiplied by -1 to facilitate comparison with Figure 3.6.

observations. Second, the subtropical SST anomaly centred on 30°N June through September has the opposite sign to observations (note that the SCA index is multiplied by -1). Third, the ECHAM4 anomaly magnitudes in the JJASO mean are ~30% lower than observed, which reflects the reduced persistence between months compared to observations. These results show that SST anomalies persisting for two or three months occur in the North Atlantic in ECHAM4 without an associated significant atmospheric signal above the ocean. However, the pattern of SST anomalies is different to that in observations associated with ΔT (Figure 3.6). This suggests that ΔT does indeed play a major role in observed NAO_{DJF} predictability through its association with North Atlantic SST.

The evidence presented above shows that the link between NH SCA and NAO_{DJF} functions differently in ECHAM4 compared to observations. First, the zonal gradients in subpolar air temperature (ΔT) associated with NH SCA are not established in ECHAM4. This is because a teleconnected contemporaneous anticyclonic anomaly south of Greenland does not form during JJA. This means that for high NH SCA conditions cooler northerly air is advected over southern Greenland, compared to warm advection from the Gulf Stream region seen in observations. Second, the absence of a ΔT link to NH SCA means that there is no significant contemporaneous response to NH SCA in zonal wind and MSLP over the North Atlantic. Third, in the absence of this atmospheric forcing in ECHAM4, SST anomalies near Newfoundland and opposite-signed subtropical SST anomalies do not evolve in the months following June. Fourth, the reduced meridional SST gradients near Newfoundland lessen the influence of North Atlantic SST on extratropical cyclogenesis and, consequently, NAO_{DJF} predictability. What remains to be proven is whether the significant link between ECHAM4 summer NH SCA and NAO_{DJF} is a physical relationship or a Type I error.

5.6 ECHAM4 stationarity 1904-2002

We now examine the stationarity in both the NAO_{DJF} predictive link and the physical mechanism discussed above. This is assessed by subdividing the 1904-2002

period into three independent segments. These segments cover the periods 1904-36, 1937-69 and 1972-2002, respectively, and are chosen because they contain an approximately equal number of years (the 1972-2002 period is retained to allow comparison with the previous analyses).

5.6.1 NAO_{DJF} predictability

The stationarity of general NAO predictability and NAO_{DJF} predictability from lagged NH SCA is assessed. The lead/lag correlation matrices between ECHAM4 monthly mean NH SCA and the NAO for 1904-36 and 1937-69 exhibit a similar level of intermonth variability as for 1972-2002 and are not field significant (not shown). This means that the general level of NAO predictability from NH SCA has remained low in ECHAM4 during the twentieth century. The stationarity in the lagged bi-monthly NH SCA/ NAO_{DJF} link is found by comparing the three panels in Figure 5.4. The significant peak linking JA SCA and NAO_{DJF} is only present for the Hurrell index 1937-69 and not for any index 1904-36. In contrast to the 1972-2002 period, the earlier plots show a significant contemporaneous link between January-February mean NH SCA and NAO_{DJF} . This represents the strong impact of the NAO circulation on snow cover, which is seen in observations 1972-2002.

The stationarity of the predictive link between JA NH SCA and NAO_{DJF} is also assessed for the earlier time periods. The top row of Table 5.1 shows the cross-validated NAO_{DJF} hindcast skill achieved 1937-69 using JA NH SCA as a lagged predictor. Positive skill is found against all three NAO_{DJF} indices, whereas 1972-2002 skill is only seen against the CRU and MSLP PC1 NAO_{DJF} indices. This is surprising because the link between JA SCA and the CRU and MSLP PC1 NAO_{DJF} indices is weaker 1937-69 than 1972-2002 (Figure 5.4). However, none of these hindcast skill values is significant using our resampling test. No positive skill was found using any other summer averaging period or for the earlier period 1904-36 (not shown). We conclude that NAO_{DJF} predictability from summer NH SCA is low in ECHAM4 throughout the twentieth century.

Prior to 1950, observations show reduced NAO_{DJF} predictability (Table 3.3).

However, it is unclear whether ECHAM4 exhibits reduced NAO_{DJF} predictability over this period for the same reasons as observations. The lack of ECHAM4 NAO_{DJF} skill could be a representation of the observed reduction in boundary forcing on NAO_{DJF} or of the dominance of internal dynamics in the CGCM. The lack of ECHAM4 NAO_{DJF} predictability prior to 1950 is somewhat surprising given the significant decadal NAO_{DJF} variability seen during the same period (Figure 5.2). However, analysis of the wavelet power spectrum for JA NH SCA reveals no low-frequency variability prior to 1950 (not shown). Figure 4.4 shows strong, but not significant, decadal variability in JJA NH SCA after 1980. Therefore, the absence of ECHAM4 NAO_{DJF} predictability 1900-50 can be attributed to a lack of decadal SCA variability, while 1972-2002 it is due to a lack of decadal NAO_{DJF} variability.

5.6.2 Physical basis for NAO_{DJF} predictability

The findings presented in Section 5.5 suggest that ECHAM4 NAO_{DJF} predictability 1972-2002 does not occur through the same mechanism as the link proposed from observations. Since NAO_{DJF} predictability from NH SCA is nonstationary 1904-2002, it is reasonable to expect that the links between NH SCA and the North Atlantic would also be nonstationary over this period. This is assessed first through the association between summer NH SCA and 2 m air temperature. Figure 5.10 shows the spatial correlation between JJA mean NH SCA and 2 m temperature poleward of 30°N for 1904-36 and 1937-69, which can be compared with Figure 5.5 for 1972-2002. Overall, the earlier patterns show strong similarities to the 1972-2002 pattern, with negative correlation above the snow covered areas and the absence of a teleconnected positive correlation south of Greenland. The teleconnection to the North Pacific is stronger in the earlier plots than 1972-2002, suggesting a greater degree of hemispheric-scale variability in this region prior to 1972. These results show that in ECHAM4 the NH SCA teleconnection to the North Atlantic is not present at any time during the twentieth century, which explains the low NAO_{DJF} predictability prior to 1950 discussed above.

Since the mechanism proposed from observations is not present in ECHAM4,

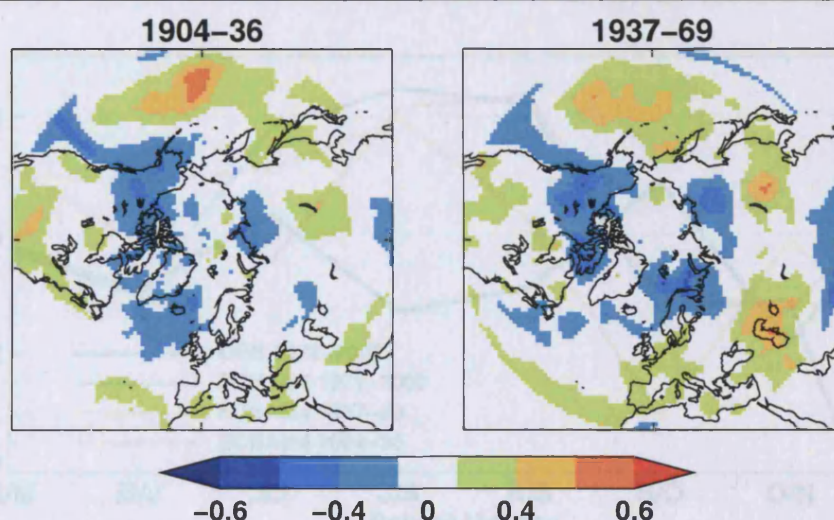


Figure 5.10: As Figure 5.5 except for ECHAM4 1904-36 (left) and 1937-69 (right).

the link between JA NH SCA and NAO_{DJF} must function through an alternative mechanism. Cohen and Saito (2003) proposed that persistence of cool conditions from summer into winter underpins the relationship between observed NH SCA and the NAO_{DJF} . This is tested in ECHAM4 by calculating the persistence in 2 m air temperature poleward of 50°N in the months before, during and following high minus low JJA NH SCA terciles. We present these data for both observations and ECHAM4 1972-2002 and for ECHAM4 during the earlier time periods 1904-36 and 1937-69. If persistence of cooler conditions is related to NAO_{DJF} predictability, then persistence should be observed in observations and ECHAM4 1972-2002 but not in ECHAM4 1904-36 when there is zero hindcast skill.

Figure 5.11 shows the month-to-month pattern correlation between fields of 2 m air temperature for pairs of months from April/May (A/M) to November/December (D/M). The month pairs J/J and J/A are contemporaneous with the NH SCA anomalies. For the period 1972-2002, persistence is low in observations ($r < 0.5$ for all months), while in ECHAM4 persistence peaks during J/A and A/S. However, the ECHAM4 pattern correlation drops to below $r = 0.3$ during S/O, which means that persistence does not extend for more than ~ 1 -2 months following JJA SCA anomalies. Furthermore, the period 1904-36 in ECHAM4 (found to have zero hindcast skill) displays a similar level of persistence as 1972-2002 from J/A to S/O.

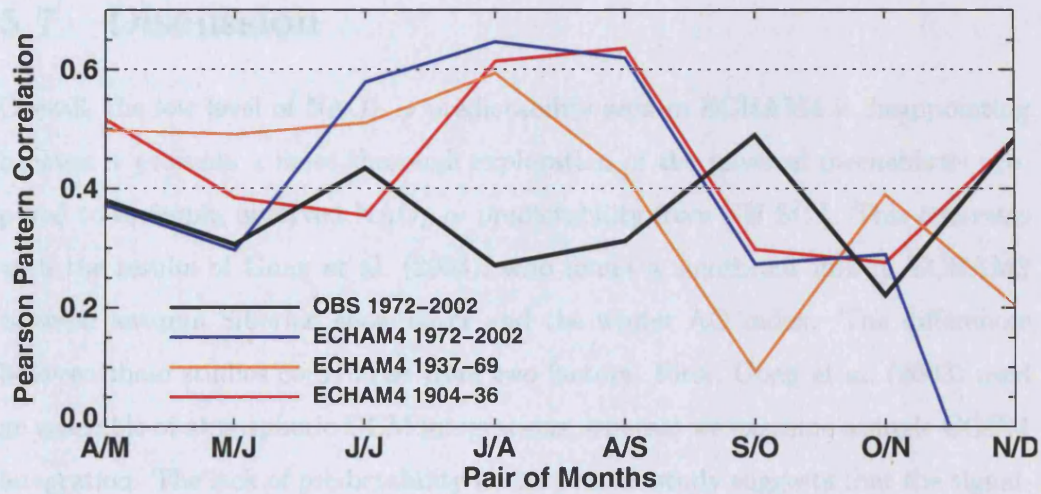


Figure 5.11: Pearson pattern correlation coefficients between fields of 2 m air temperature for pairs of adjacent months denoted by their first letters. Temperature fields are the composite mean associated with high minus low JJA NH SCA terciles. Line colours denote: observations 1972-2002 (black), ECHAM4 1972-2002 (blue), ECHAM4 1937-69 (orange) and ECHAM4 1904-36 (red). Dashed horizontal lines show pattern correlations at $r = 0.4$ and $r = 0.6$, respectively.

These results suggest that persistence in air temperature is unrelated to NH SCA anomalies (and consequently NAO_{DJF} predictability) both in observations and in ECHAM4.

There are two main conclusions from this analysis. First, during the period when ECHAM4 NH SCA is related significantly to NAO_{DJF} (1972-2002), the link does not occur through the same mechanism as seen in observations. Furthermore, the sign, location and magnitude of the NH SCA-related North Atlantic SST anomalies make them unlikely to influence the NAO_{DJF} directly. Second, ECHAM4 NAO_{DJF} predictability from NH SCA is nonstationary 1904-2002. However, periods of high NAO_{DJF} predictability are unrelated to persistence in cooler air temperatures from summer to winter. Therefore, we conclude that the significant link seen in ECHAM4 between JA NH SCA and NAO_{DJF} is most likely the result of a Type I error, rather than a physical relationship.

5.7 Discussion

Overall, the low level of NAO_{DJF} predictability seen in ECHAM4 is disappointing because it prevents a more thorough exploration of the physical mechanisms proposed to underpin observed NAO_{DJF} predictability from NH SCA. This contrasts with the results of Gong et al. (2003), who found a significant link in ECHAM3 between autumn Siberian snow cover and the winter AO index. The differences between these studies could arise from two factors. First, Gong et al. (2003) used an ensemble of atmospheric GCM integrations, whereas we examine a single CGCM integration. The lack of predictability in the present study suggests that the signal-to-noise ratio from a single ECHAM4 integration is not sufficient to detect contemporaneous or lagged snow/ NAO_{DJF} links. This means the ability to distinguish snow-related climate responses from internal variability is restricted. Second, the transient coupled ECHAM4 integration is subject to climatological drift. Moreover, this study has shown that coupling the atmosphere and ocean in ECHAM4 may result in less accurate boundary forcing of the atmosphere than using prescribed ocean conditions. This is because oceanic variability may diverge from observations during the transient integration 1860-2050. Further investigation is required into the influence of coupling on the boundary forcing.

An important finding highlighted in Section 5.5 is that the ΔT link to the North Atlantic sector depends more on the teleconnection with southern Greenland than with the snow covered regions of EU and NA. The key dynamical property of the teleconnection appears to be an anticyclonic anomaly south of Greenland associated with high NH SCA, which advects warm air polewards from the Gulf Stream region during summer. These findings are in agreement with the discussion in Chapter 3, which stated that in observations NH snow cover explains only $\sim 50\%$ of the variance in summer ΔT . In ECHAM4, the ΔT variance explained is less than 10% with Greenland included and $\sim 20\%$ excluding Greenland. Therefore, the inaccurate representation of teleconnections in ECHAM4 means summer snow cover is less related to ΔT (and to NAO_{DJF}) than in observations.

The air temperature response above Greenland in ECHAM4 is accurate during

winter despite the absence of snow cover on the Greenland ice sheet. However, larger errors were found over Greenland during summer, which raises several possibilities. First, the atmospheric response to the lack of Greenland snow could be seasonal. Second, the observed seesaw pattern is weaker during summer and its hemispheric-scale effects are reduced. Therefore, ECHAM4 may capture the reduced NH climate variability during summer but with large errors in the teleconnected temperature responses. Third, the model may have been tuned to perform well during winter, when climate variability is increased and predictability is greater. From the analysis in this study, it is not apparent which of these explanations is most likely.

The observational evidence presented in this thesis supports an active, dynamical teleconnection between NH SCA and the North Atlantic sector. The results found in ECHAM4 are therefore contradictory to the observational results. This study was intended as a preliminary investigation to highlight the potential use of CGCMs in investigating lagged NAO_{DJF} predictability from NH snow cover. The next stage will be to employ an ensemble of CGCM integrations forced with realistic NH SCA scenarios to determine the precise role of NH SCA in NAO_{DJF} predictability. This should yield more information on the potential for seasonal NAO_{DJF} predictability from lagged snow cover.

5.8 Limitations of CGCMs for snow/atmosphere modelling

This study has highlighted the strengths and weaknesses of using a twentieth century CGCM integration for snow/atmosphere investigations. One of the key aims of this research was to evaluate the limitations in CGCM snow modelling and to propose potential improvements that could be made to future CGCMs for this purpose. Four main limitations were highlighted in the ECHAM4 CGCM, three of which are likely to affect other CGCMs.

First, the boundary forcing and response of hemispheric and regional SCA indices was different in ECHAM4 and observations. Evidence to support this assertion

comes from a variety of sources. The correlation between the time series of observed and ECHAM4 interannual SCA variability 1972-2002 is low ($r_{RANK} < 0.3$) for all seasons except summer, when ECHAM4 underestimates the magnitude of SCA variability by $\sim 60\%$. The presence of multi-annual persistence and decadal SCA variability is not captured by ECHAM4. Furthermore, trends in ECHAM4 SCA exhibit less seasonality compared to observations in response to a realistic increase in mean air temperature 1972-2002. Improvements to schemes coupling the land, ocean and atmosphere are essential to ensure that the time scales of variability from snow to atmosphere and atmosphere to snow are represented accurately. Such improvements are already incorporated into the latest generation of CGCMs and future validation studies will show what benefits this has brought.

Second, climate variability in ECHAM4 appears to occur on more regional spatial scales than in observations. This is demonstrated in analyses of ECHAM4 SCA, NAO and NH geopotential height variability. One explanation for this could be the relative dominance of internal (and therefore smaller scale), compared to boundary-forced (larger scale), atmospheric variability in the CGCM. This is highlighted by a lack of low-frequency (i.e., large scale) NAO or SCA variability 1904-2002 and by the absence of a contemporaneous link between NH SCA and NAO_{DJF}. The effect of this limitation is that hemispheric-scale teleconnections and NAO_{DJF} predictability appear to be lower in ECHAM4 than in observations. These limitations may be reduced in an ensemble integration, which would filter the internal model variability.

Third, the correlation between observed and ECHAM4 seasonal snow-related surface fluxes is $r_{RANK} \sim -0.15$ for EU and $r_{RANK} \sim 0.30$ for NA 1972-2002. Since the contact between the snow surface and the atmosphere occurs through these fluxes, this limitation means ECHAM4 would not be expected to capture correctly the *in situ* or teleconnected atmospheric responses to snow cover. Yet, the *in situ* atmospheric temperature responses to regional SCA over NA and EU are modelled accurately in ECHAM4. Further investigation is required to determine the reason for this apparent contradiction. The CGCM may be deficient in the representation of cloud cover above the snow surface and/or soil moisture processes related to snow

melt. These factors are difficult to verify because of the lack of reliable observational data for comparisons.

Fourth, the ECHAM4 GSDIO integration was run with no snow cover on the Greenland ice sheet. An explanation for this limitation was not forthcoming from MPI or DKRZ despite repeated requests. Therefore, it is unknown whether the effect of Greenland snow was parameterised in GSDIO through another variable, for example, glacier thickness or ice depth. One of the problems in the ECHAM4 representation of NH SCA/ NAO_{DJF} links was found to be the representation of the summer air temperature teleconnection between EU, NA and southern Greenland. Although tests showed that the air temperature response above Greenland is accurate during winter, greater errors are seen in this region during summer. Therefore, the absence of snow cover over Greenland may contribute, at least in part, to the reduced NAO_{DJF} predictability from summer snow cover in ECHAM4.

5.9 Summary and conclusions

This Chapter presents an analysis of twentieth century ECHAM4 NAO_{DJF} predictability associated with lagged NH snow cover. ECHAM4 NAO_{DJF} variability corresponds poorly to observations 1904-2002. The general level of NAO predictability available from lagged NH SCA 1972-2002 is low in observations and ECHAM4. Furthermore, ECHAM4 does not capture the contemporaneous NAO_{DJF} influence on SCA. However, a significant link is identified between July-August (JA) mean NH SCA and NAO_{DJF} , which produces positive, but not statistically significant, hindcast skill.

The physical mechanism linking ECHAM4 summer NH SCA and NAO_{DJF} functions differently to the link in observations. The contemporaneous zonal gradients in summer subpolar air temperature (ΔT), associated with NH snow cover in observations, are not established in ECHAM4. This is because the ECHAM4 response to NH SCA is more regional in scale. Consequently, the observed teleconnected anticyclonic anomaly over the North Atlantic, associated with high NH SCA, is not found during summer. The contemporaneous atmospheric and lagged oceanic

responses to NH SCA over the North Atlantic do not correspond to observations, which means the SST feedback onto NAO_{DJF} during autumn is absent. The JA NH SCA/ NAO_{DJF} link in ECHAM4 is nonstationary during the twentieth century, which is not explained by ΔT or by persistence in cooler temperatures from summer into winter. It is therefore likely that the link in ECHAM4 is a Type I error rather than a physical association.

Four main limitations for snow/atmosphere modelling are identified in ECHAM4. First, the boundary forcing and response of SCA is different in observations and ECHAM4. Second, compared to observations, climate variability in ECHAM4 occurs on more regional spatial scales. This is most likely related to the dominance of internal dynamics in ECHAM4. Third, problems were identified in the energy balance of snow covered surfaces, which may result from errors in cloud parameterisation or deficiencies in soil moisture hydrology. Fourth, the ECHAM4 run contained no snow over the Greenland ice sheet, which may lower NAO_{DJF} predictability during summer.

Chapter 6

Conclusions and Future Work

6.1 Introduction

This thesis has investigated the predictability of North Atlantic winter climate on seasonal-to-interannual time scales. Chapter 1 frames the problem of seasonal climate forecasting (SCF) and discusses two alternative methodologies for SCF. The motivation for the specific areas of research in this thesis is also presented. Climate data requirements for these research areas are discussed in Chapter 2. The results from the original research are reported in Chapters 3, 4 and 5. Chapter 3 examines empirical predictability for the winter NAO (NAO_{DJF}) and compares four published prediction schemes using a standardised assessment. Chapter 4 assesses the capability of a coupled GCM (CGCM) to represent the observed temporal and spatial variability in Northern Hemisphere (NH) snow cover 1972-2002. Chapter 5 quantifies the NAO_{DJF} predictability associated with NH summer snow cover in the CGCM simulation and investigates its stationarity during the twentieth century. This Chapter summarises the main conclusions from this work and provides answers to the research questions posed in Sections 1.4.6 and 1.5.4. Potential future research arising from this work is also discussed.

6.2 Best lagged NAO_{DJF} predictor (Chapter 3)

Chapter 3 presents a detailed assessment of the current level of seasonal empirical NAO_{DJF} predictability. A standardised hindcast procedure is used to validate four previously published lagged NAO_{DJF} predictors. A new predictor based on the

zonal gradient in summer NH subpolar 2 m air temperature (ΔT) is also examined.

(i) *What is the most skilful lagged predictor of the NAO_{DJF} ?*

Over three extended assessment periods out to 100-years, summer ΔT is most skilful in predicting NAO_{DJF} . For the period 1900-2001, May-September mean ΔT offers highest skill ($\sim 6\text{-}9\%$ improvement over climatology). Since 1972, June-July mean ΔT produces highest skill ($\sim 35\%$ improvement over climatology).

(ii) *Are the predictive NAO_{DJF} relationships stationary when assessed over 100, 50 and 30 year periods?*

An increase in NAO_{DJF} predictability since 1972 is observed for all predictors except those derived using SSTs. Our findings concur with those of previous studies (e.g., Rodwell and Folland 2002, Trigo et al. 2004), who found that predictability has increased since 1950. Skill from the summer ΔT predictor exhibits a ~ 60 year cycle. Predictability from SSTs is highly variable since 1950 and depends on the specific regions, time periods and statistical methods employed. It is unclear whether the nonstationarity is a result of increasing greenhouse gas emissions or a natural feature of the climate system (Trigo et al. 2002).

(iii) *Does the most skilful period coincide with other periods of variability in either NAO_{DJF} or the predictors?*

The increase in skill since 1972 coincides with a period of high decadal NAO_{DJF} variability. The intervals of highest predictability from the $\Delta T/NAO_{DJF}$ link coincide with positive trends in the NAO_{DJF} index in the early and late twentieth century. However, ΔT performs equally well predicting above or below median NAO_{DJF} seasons.

(iv) *What are the physical mechanisms that underpin the link between the NAO_{DJF} and the most skilful lagged predictor?*

Evidence is presented supporting a physical link between summer ΔT and NAO_{DJF} . NH snow cover plays a significant role in establishing ΔT in summer. During subsequent months the atmospheric circulation response to ΔT is centred on the

midlatitude North Atlantic. These circulation anomalies lead by ~ 1 month the formation of a pattern of North Atlantic SSTs, which persist into late autumn. SST persistence is strong off southeast Newfoundland, which coincides with the main region of North Atlantic cyclogenesis. Therefore, through meridional SST gradients and *in situ* surface fluxes, the SST pattern could influence the formation of extratropical cyclones and, subsequently, the NAO_{DJF} .

Further research

This study has focused on predicting NAO_{DJF} . Future investigations should determine the predictability found for NAO indices computed over different averaging periods. For example, Thompson and Wallace (2000) found that the North Atlantic climate is most active January through March. Linear statistical techniques are employed to model the relationships between the lagged predictors and the NAO_{DJF} . However, there is no reason to suggest that these links are best described using linear models. Several studies have highlighted asymmetrical relationships between the climate system and positive and negative NAO_{DJF} phases (e.g., Trigo et al. 2002, Peng et al. 2003). Future investigations employing nonlinear regression techniques (e.g., Dewar and Wallis 1999) could be conducted to assess the degree of nonlinearity in the predictive relationships.

NAO_{DJF} predictability may be further improved by clarifying the causes of nonstationarity during the twentieth century. For example, skill from persistent North Atlantic SSTs during autumn is higher in some years than others. Future research could employ observational data or an atmospheric GCM ensemble to investigate why these links are nonstationary. Furthermore, NAO_{DJF} predictability arising from North Atlantic SSTs has been shown to compete with the influences of ENSO and Indian Ocean SST variability (Hoerling et al. 2001, Sutton and Hodson 2003). Future research should aim to quantify the exact magnitude of the North Atlantic influence by taking steps to isolate the North Atlantic signal from that of the other basins. In observational studies this could be approximated by analysing the residuals from a regression of ENSO on North Atlantic SST. However, the

impact on the atmosphere from SSTs in one particular region is best examined using a GCM forced with fixed or climatological SSTs except in the region of interest.

6.3 Coupled GCM snow validation (Chapter 4)

Chapter 4 presents a comparison of the temporal and spatial variability of seasonal snow covered area (SCA) in observations and a coupled GCM (CGCM). The CGCM employed is the ECHAM4/OPYC3 scheme (henceforth ‘ECHAM4’) sponsored jointly by the Max Planck Institut für Meteorologie and the Deutsches Klimarechenzentrum.

- (i) *Can CGCM snow depth data be employed to create an index of monthly mean snow cover?*

Indices of fractional SCA 1904-2002 were computed from snow depth output taken from the ECHAM4 twentieth century coupled climate integration. The SCA indices represent monthly mean snow cover over North America (NA), Eurasia (EU) and the entire NH poleward of 20°N. Grid cells are considered to be completely snow covered above a threshold depth of 2.5 cm. Greenland was shown to have zero snow depth lying on the ice sheet throughout the ECHAM4 integration. The NA and NH SCA indices were therefore corrected to include the Greenland land area.

- (ii) *How well does a coupled GCM simulate the observed spatial and temporal variability in seasonal snow cover when forced with observed radiative forcing 1972-2002?*

Negative linear trends in ECHAM4 SCA are stronger than observations in winter and weaker in summer. These trends occur in response to a realistic increase in mean air temperature, which may cause overestimation of the winter snow/albedo feedback in ECHAM4. The annual SCA cycle in ECHAM4 compares closely with observations. However, seasonally varying regional biases are found, which are largest over NA during summer and over EU during spring. The representation of observed interannual SCA variability is poor for all regions and all seasons. Compared to observations, ECHAM4 summer SCA shows highest interannual cor-

relations but only one-third of the variability. Furthermore, observed multi-annual persistence of winter and summer SCA anomalies is not captured by ECHAM4. This evidence supports the theory that ECHAM4 SCA is driven primarily by internal dynamics and that the boundary forcing and response of ECHAM4 SCA are different to observations.

- (iii) *Does the coupled GCM simulate realistically the observed contemporaneous in situ links between seasonal snow cover and the atmosphere 1972-2002?*

High seasonal SCA is associated with an equivalent barotropic cooling, which extends from the surface to the tropopause above the snow covered regions. An opposite-signed anomaly is seen in the stratosphere. ECHAM4 reproduces these observed patterns with accuracy in spatial position and magnitude, although differences are most pronounced during winter and autumn in both EU and NA. ECHAM4 captures accurately the significant lagged snow/temperature feedback in summer. However, ECHAM4 seasonal mean surface fluxes are accurate only over NA during winter and spring. The flux response above snow cover shows the largest errors in ECHAM4, particularly during summer. Furthermore, the observed contemporaneous link between winter NH SCA and NAO_{DJF} is not captured in ECHAM4.

- (iv) *Does coupling a GCM to the ocean improve its representation of snow cover variability?*

Compared to atmospheric GCMs, ECHAM4 shows some improvements in the representation of the annual SCA cycle and *in situ* atmospheric temperature responses to SCA (Frei and Robinson 1998, Frei et al. 2003). However, regional biases and the dominance of internal dynamics in SCA variability are not improved in the coupled model. Twentieth century boundary forcing of SCA and the NAO_{DJF} are most likely different to observations. Furthermore, serious limitations are found in the surface energy balance above snow cover. Improvements are expected from the next generation of CGCMs, most of which are currently operational without flux adjustments (Covey et al. 2003).

Further research

The validation of GCMs will always be limited by the lack of a definitive set of observational data with which to compare the models. In this study, NCEP/NCAR reanalysis data were employed for temperature, winds and surface fluxes but these data are themselves derived from a CGCM and are not without inhomogeneities. To minimise these effects, future comparisons of observed and CGCM snow/atmosphere interaction should be performed using as wide a selection of observational/reanalysis data as possible, e.g., ERA-40 data. Where possible, this should include station observations.

The focus on the ECHAM4 CGCM allowed for a more detailed examination of snow/atmosphere interaction than in previous CGCM snow validations (e.g., Frei and Robinson 1998, Frei et al. 2003). However, future research investigating NAO_{DJF} predictability should seek to minimise the effect of individual CGCM errors by comparing results from a suite of CGCMs. The deficiencies in the ECHAM4 surface fluxes and albedo may be better simulated in other models. For example, data are available from a twentieth century integration run on the UKMO coupled HadCM3 model. However, Covey et al. (2003) found that current CGCMs share many of the same deficiencies. Therefore, improvements may not be seen until the next generation of coupled GCMs become available to the research community.

Some of the inaccuracies identified in the CGCM response to SCA may be due to the formulation of the SCA indices. The Frei et al. (2003) method is not ideal because all grid cells with snow depth > 2.5 cm are assigned the same weight. More complex SCA indices could be computed and these may yield improvements in apparent CGCM performance. One example of an alternative SCA index is to use the leading principal component of snow depth. However, SCA indices involving snow depth have the limitation that they diverge from the observational snow cover indices, which hinders comparison. Recent research also suggests that increased information on snow/atmosphere interaction can be obtained by weighting snow cover indices by latitude and insolation (Pielke et al. 2004).

6.4 NAO_{DJF} predictability from snow cover in a coupled GCM (Chapter 5)

Chapter 5 presents an analysis of twentieth century ECHAM4 NAO_{DJF} predictability associated with lagged NH snow cover.

- (i) *Does a coupled GCM capture observed temporal NAO_{DJF} variability during the twentieth century?*

The range and magnitude of NAO_{DJF} variability is captured accurately. However, three ECHAM4 NAO_{DJF} indices exhibit several differences to three equivalent observational NAO_{DJF} indices. First, the CGCM and observed indices show low correspondence 1904-2002. Second, there is little trend in the ECHAM4 indices 1904-2002 or 1972-2002. Third, ECHAM4 fails to capture correctly the timing of observed peaks in decadal NAO_{DJF} variability. These factors contribute to lower NAO_{DJF} predictability in ECHAM4 than in observations.

- (ii) *Does a coupled GCM represent accurately the observed lagged and contemporaneous links between snow cover and the NAO_{DJF} 1972-2002?*

The general level of NAO predictability available from lagged NH SCA 1972-2002 is low in observations and ECHAM4. ECHAM4 does not capture the contemporaneous NAO_{DJF} influence on SCA, however, a significant link is identified between July-August (JA) mean NH SCA and NAO_{DJF} . This is one month later than the link identified in observations (Saunders et al. 2003). Cross-validated hindcasts using ECHAM4 JA NH SCA as a lagged NAO_{DJF} predictor show some positive, but not statistically significant, skill.

- (iii) *Is the physical mechanism linking summer snow cover and the NAO_{DJF} the same in the coupled GCM as in observations?*

The physical mechanism linking JA NH SCA and NAO_{DJF} in ECHAM4 functions differently to the link proposed from observations. The contemporaneous zonal gradients in summer subpolar air temperature (ΔT), associated with NH snow cover in observations, are not established in ECHAM4. This is because the ECHAM4

teleconnected response to NH SCA is more regional in scale than in observations. Consequently, the contemporaneous teleconnected anticyclonic anomaly over the North Atlantic, associated with high NH SCA, does not form during summer. During the months following JA, persistent North Atlantic SST anomalies evolve in ECHAM4 but they exhibit different magnitudes and spatial positions compared to observations. Therefore, the feedback to the atmosphere from meridional SST gradients in the region near Newfoundland does not occur (*c.f.* Section 3.4.4).

- (iv) *Is the link between summer snow cover and the NAO_{DJF} stationary during the twentieth century?*

The significant link between JA NH SCA and NAO_{DJF} is present during the periods 1937-69 and 1972-2002 but not during 1904-36. However, neither summer ΔT nor persistence in air temperatures explain this significant link. Persistence of cooler conditions is seen for 1-2 months following summer NH SCA anomalies. However, persistence is also seen during the 1904-36 period, when there is no significant link between JA NH SCA and NAO_{DJF} . It is therefore likely that the link in ECHAM4 is a Type I error rather than a physical association.

- (v) *What are the main limitations in using a coupled GCM for snow/atmosphere modelling?*

Four main limitations are identified in ECHAM4. First, the boundary forcing and response of the ECHAM4 SCA indices is different to observations. This suggests a need for improved CGCM land surface schemes, which should be available from the next generation of coupled models. Second, compared to observations, climate variability in ECHAM4 occurs on more regional spatial scales. This is most likely related to the dominance of internal dynamics in ECHAM4, which acts on smaller spatial scales than boundary-forced variability (e.g., Lorenz 1969). These limitations would be reduced in an ensemble integration. Third, problems were identified in the radiative balance of snow covered surfaces, which may result from errors in cloud parameterisation or deficiencies in soil moisture hydrology. Fourth, the ECHAM4 run contained no snow over the Greenland ice sheet. During winter the

atmospheric response over Greenland was unaffected. However, larger differences to observations were found over this region during summer.

Further research

Further investigation is required to clarify the physics underpinning the mechanism linking summer NH snow cover and the NAO_{DJF} . The relationship between snow cover, ΔT and the teleconnection to the North Atlantic could be examined using CGCM experiments forced with observed and/or idealised snow cover scenarios. However, this would require a CGCM capable of responding on hemispheric, rather than purely regional, spatial scales. Another solution would be to analyse the output fields from the DEMETER project (Palmer et al. 2004) to assess whether the link is better captured in the multi-model ensemble. The influence of the summer zonal wind anomalies on North Atlantic SST near Newfoundland also needs to be quantified. One simple method would be to examine the SST evolution in response to an idealised wind field imposed on a passive ocean. The NAO_{DJF} response in the idealised experiment could then be compared to the NAO_{DJF} response to climatological SSTs.

Snow cover and soil moisture are related significantly, particularly during spring and summer (Ose 1996, Shinoda 2001, Robock et al. 2003). Some of the limitations identified in ECHAM4 snow/atmosphere interaction may result from deficiencies in the CGCM hydrological cycle. A comparison between observed and modelled soil moisture fields should reveal how well the CGCM simulates hydrological snow feedbacks. However, a major limitation is that long records of soil moisture are restricted to station observations over China and the Former Soviet Union (Robock et al. 2003).

This study has focused on the predictability of the NAO_{DJF} , which is the dominant mode of variability for North Atlantic winter climate. However, the NAO_{DJF} mode explains $\sim 40\%$ of the variance in twentieth century winter North Atlantic MSLP (Hurrell et al. 2003), which means other modes explain the remaining 60%. Barnston and Livezey (1987) described several such modes, including the East At-

lantic, Eurasia-1 (or Scandinavian) and Eurasia-2 (or East Atlantic/West Russia) patterns. Future work should assess whether additional North Atlantic predictability can be derived from the influence of lagged snow cover on these patterns of variability.

6.5 Final remarks

NAO_{DJF} predictability above the levels reported previously for the period 1972-2002 has been identified from empirical models. However, nonstationarity in the predictive relationships means that, to a large extent, we cannot rely on these models for future NAO_{DJF} predictions. CGCMs offer significant opportunities for resolving these issues through detailed investigation of the physics underpinning the empirical relationships. Yet this work has highlighted major problems in the representation of climate processes in current CGCMs. Future generations of CGCMs may be more accurate, but experience has shown that it is not easy to make big improvements in these models (van Oldenborgh et al. 2003). The way forward for seasonal forecasting must be increasing interaction between empirical and dynamical modellers and a realisation that both camps depend on each other for future progress.

References

- Ambaum, M. H. P., B. J. Hoskins, and D. B. Stephenson, 2001: Arctic Oscillation or North Atlantic Oscillation? *J. Climate*, **14**, 3495–3506.
- Baldwin, M. P. and T. P. Dunkerton, 2001: Stratospheric harbingers of anomalous weather regimes. *Science*, **294**, 581–584.
- Bamzai, A. S., 2003: Relationship between snow cover variability and Arctic Oscillation index on a hierarchy of time scales. *Int. J. Climat.*, **23**, 131–142.
- Barnston, A. G., 1994: Linear statistical short-term climate predictive skill in the Northern Hemisphere. *J. Climate*, **7**, 1513–1564.
- Barnston, A. G. and R. E. Livezey, 1987: Classification of seasonality and persistence of low-frequency atmospheric circulation patterns. *Mon. Wea. Rev.*, **115**, 1083–1126.
- Barnston, A. G., S. J. Mason, L. Goddard, D. G. DeWitt, and S. E. Zebiak, 2003: Multimodel ensembling in seasonal climate forecasting at IRI. *Bull. Amer. Meteorol. Soc.*, **84**, 1783–1796.
- Basnett, T. and D. Parker, 1997: Development of the Global Mean Sea Level Pressure data set GMSLP2. Climate research technical note 79, Hadley Centre, Met Office, Fitzroy Road, Exeter, Devon, EX1 3PB, UK.
- Becker, B. D., J. M. Slingo, L. Ferranti, and F. Molteni, 2001: Seasonal predictability of the Indian Summer Monsoon: What role do land surface conditions play? *MAUSAM*, **52**, 175–190.
- Bernstein, R. L. and D. B. Chelton, 1985: Large-scale sea surface temperature variability. *J. Geophys. Res.*, **90**, 11619–11630.
- Bjerknes, J., 1964: Atlantic air-sea interaction. *Advances in Geophysics*, **10**, 1–82.
- Bjerknes, J., 1966: A possible response of the atmospheric Hadley circulation to equatorial anomalies of ocean temperature. *Tellus*, **18**, 820–829.
- Blandford, H. F., 1884: On the connection of the Himalaya snowfall with dry winds and seasons of drought in India. *Proc. Roy. Soc.*, **37**, 3–22.
- Bloomfield, P., 2000: *Fourier Analysis of Time Series*. John Wiley and Sons, New York, 288pp.

References

- Bojariu, R. and L. Gimeno, 2003: The role of snow cover fluctuations in multiannual NAO persistence. *Geophys. Res. Lett.*, **30**, doi:10.1029/2002GL015651.
- Bottomley, M., C. K. Folland, J. Hsiung, R. E. Newell, and D. E. Parker, 1990: *Global ocean surface temperature atlas*. HMSO, London.
- Branković, C. and T. N. Palmer, 2000: Seasonal skill and predictability of ECMWF PROVOST ensembles. *Quart. J. Roy. Meteor. Soc.*, **126**, 2035–2067.
- Bretherton, C. S. and D. S. Battisti, 2000: An interpretation of the results from atmospheric general circulation models forced by the time history of the observed sea surface temperature distribution. *Geophys. Res. Lett.*, **27**, 767–770.
- Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. *Mon. Wea. Rev.*, **75**, 1–3.
- Brock, F. W. and S. J. Richardson, 2001: *Meteorological Measurement Systems*. Oxford University Press, New York, 290pp.
- Brown, R. D., 2000: Northern Hemisphere snow cover variability and change, 1915–97. *J. Climate*, **13**, 2339–2355.
- Cassou, C., C. Deser, L. Terray, J. W. Hurrell, and M. Drévillon, 2004: Summer sea surface temperature conditions in the North Atlantic and their impact upon the atmospheric circulation in early winter. *J. Climate*, **17**, 3349–3363.
- Chen, C. T. and E. Roeckner, 1996: Validation of the earth radiation budget as simulated by the Max Planck Institute for Meteorology general circulation model ECHAM4 using satellite observations of the earth radiation budget experiment. *J. Geophys. Res.*, **101**, 4269–4287.
- Clark, M. P., M. P. Serreze, and D. A. Robinson, 1999: Atmospheric controls on Eurasian snow extent. *Int. J. Climat.*, **19**, 27–40.
- Claussen, M., U. Lohmann, E. Roeckner, and U. Schulzweida, 1994: *A global dataset of land surface parameters*. MPI Report No. 135. Max-Planck-Institute for Meteorology, Hamburg, Germany.
- Clemmons, L., 2002: Introduction to weather risk management. In *Weather risk management*, Re Capital Products Inc, E., editor. Palgrave, 10–25.
- Cline, D. W., 1997: Snow surface energy exchanges and snowmelt at a continental, midlatitude Alpine site. *Water Resources Research*, **33**, 689–701.

- CMIP, 2005: *Model ECHAM4+OPYC3: Elaborations*. Coupled Model Intercomparison Project. <http://www-pcmdi.llnl.gov/projects/modeldoc/cmip/>.
- Cohen, J. and D. Entekhabi, 1999: Eurasian snow cover variability and Northern Hemisphere climate predictability. *Geophys. Res. Lett.*, **26**, 345–348.
- Cohen, J. and D. Rind, 1991: The effect of snow cover on the climate. *J. Climate*, **4**, 689–706.
- Cohen, J. and K. Saito, 2003: Eurasian snow cover, more skillful in predicting U.S. winter climate than the NAO/AO? *Geophys. Res. Lett.*, **30**, doi:10.1029/2003GL018053.
- Colman, A. and M. Davey, 1999: Prediction of summer temperature, rainfall, and pressure in Europe from preceding winter North Atlantic ocean temperature. *Int. J. Climat.*, **19**, 513–536.
- Cornillon, P. and L. Strama, 1985: The distribution of diurnal sea surface warming events in the western Sargasso Sea. *J. Geophys. Res.*, **90**, 11811–11815.
- Covey, C., K. M. AchutaRao, U. Cubasch, P. Jones, S. J. Lambert, M. E. Mann, T. J. Phillips, and K. E. Taylor, 2003: An overview of the results from the Coupled Model Intercomparison Project. *Global and Planetary Change*, **37**, 103–133.
- Crook, N. A., 1996: Sensitivity of moist convection forced by boundary layer processes to low-level thermodynamic fields. *Mon. Wea. Rev.*, **124**, 1767–1785.
- Cullen, H. M. and P. B. deMencol, 2000: North Atlantic influence on Tigris-Euphrates streamflow. *Int. J. Climat.*, **20**, 853–863.
- Czaja, A. and C. Frankignoul, 2002: Observed impact of Atlantic SST anomalies on the North Atlantic Oscillation. *J. Climate*, **15**, 606–623.
- Davies, J. R., D. P. Rowell, and C. K. Folland, 1997: North Atlantic and European seasonal predictability using an ensemble of multidecadal atmospheric GCM simulations. *Int. J. Climat.*, **17**, 1263–1284.
- Davis, R. E., 1976: Predictability of sea surface temperature and sea level pressure anomalies over the north Pacific Ocean. *J. Phys. Oceanogr.*, **6**, 249–266.
- Deser, C., M. A. Alexander, and M. S. Timlin, 2003: Understanding the persistence of sea surface temperature anomalies in midlatitudes. *J. Climate*, **16**, 57–72.

- Dewer, R. E. and J. R. Wallis, 1999: Geographical patterning of interannual rainfall variability in the tropics and near tropics: An L-moments approach. *J. Climate*, **12**, 3457–3466.
- Drévilion, M., L. Terray, P. Rogel, and C. Cassou, 2001: Midlatitude SST influence on European winter climate variability in the NCEP reanalysis. *Climate Dyn.*, **18**, 331–344.
- Eden, C. and T. Jung, 2000: North Atlantic interdecadal variability: Oceanic response to the North Atlantic Oscillation (1865–1997). *J. Climate*, **14**, 676–691.
- Enfield, D. B., 1996: Relationships of inter-American rainfall to tropical Atlantic and Pacific SST variability. *Geophys. Res. Lett.*, **23**, 3505–3508.
- Etchevers, P., E. Martin, and Coauthors, 2002: SnowMIP, an intercomparison of snow models: first results. In *Proceedings of the International Snow Science Workshop, Penticon, B.C., Canada*.
- Fasullo, J., 2004: A stratified diagnosis of the Indian Monsoon-Eurasian snow cover relationship. *J. Climate*, **17**, 1110–1122.
- Feldstein, S. B., 2002: The recent trend and variance increase of the Annual Mode. *J. Climate*, **15**, 88–94.
- Folland, C. K. and D. E. Parker, 1995: Correction of instrumental biases in historical sea-surface temperature data. *Quart. J. Roy. Meteor. Soc.*, **121**, 319–367.
- Foster, J. L., G. Liston, and Coauthors, 1996: Snow cover and snow mass intercomparison of general circulation models and remotely sensed datasets. *J. Climate*, **9**, 409–426.
- Frei, A., J. A. Miller, and D. A. Robinson, 2003: Improved simulations of snow extent in the second phase of the Atmospheric Model Intercomparison Project (AMIP-2). *J. Geophys. Res.*, **108**, 4369, doi:10.1029/2002JD003030.
- Frei, A. and D. A. Robinson, 1998: Evaluation of snow extent and its variability in the Atmospheric Model Intercomparison Project. *J. Geophys. Res.*, **103**, 8859–8871.
- Fu, L., D. B. Chelton, and V. Zlotnicki, 1988: Satellite altimetry: Observing ocean variability from space. *Oceanography*, **1**, 4–11.
- Fu, Q., C. M. Johanson, S. G. Warren, and D. J. Seidel, 2004: Contribution of

- stratospheric cooling to satellite-inferred tropospheric temperature trends. *Nature*, **429**, 55–58.
- Gilchrist, A., 1986: Long-range forecasting. *Quart. J. Roy. Meteor. Soc.*, **112**, 567–592.
- Goddard, L., S. J. Mason, S. E. Zebiak, C. F. Ropelewski, R. Basher, and M. A. Cane, 2001: Current approaches to seasonal-to-interannual climate predictions. *Int. J. Climat.*, **21**, 1111–1152.
- Goldberg, M. D. and L. M. McMillin, 1999: Methodology for deriving deep-layer mean temperatures from combined satellite infrared and microwave observations. *J. Climate*, **12**, 5–20.
- Goldenberg, S. B., C. W. Landsea, A. M. Mestas-Núñez, and W. M. Gray, 2001: The recent increase in Atlantic hurricane activity: causes and implications. *Science*, **293**, 474–479.
- Gong, G., D. Entekhabi, and J. Cohen, 2002: A large-ensemble model study of the wintertime AO/NAO and the role of interannual snow perturbations. *J. Climate*, **15**, 3488–3499.
- Gong, G., D. Entekhabi, and J. Cohen, 2003: Modeled Northern Hemisphere winter climate response to realistic Siberian snow anomalies. *J. Climate*, **16**, 3917–3931.
- Gong, G., D. Entekhabi, J. Cohen, and D. Robinson, 2004: Sensitivity of atmospheric response to modeled snow anomaly characteristics. *J. Geophys. Res.*, **109**, art. no. D06107.
- Graham, R. J., A. D. L. Evans, K. R. Mylne, M. S. J. Harrison, and K. B. Robertson, 2000: An assessment of seasonal predictability using atmospheric general circulation models. *Quart. J. Roy. Meteor. Soc.*, **126**, 2211–2240.
- Groisman, P. Y., T. R. Karl, R. W. Knight, and G. L. Stenchikov, 1994: Changes of snow cover, temperature and radiative heat balance over the Northern Hemisphere. *J. Climate*, **7**, 1633–1656.
- Hanssen-Bauer, I. and E. Forland, 2001: Verification and analysis of a climate simulation of temperature and pressure fields over Norway and Svalbard. *Climate Res.*, **16**, 225–235.
- Hartmann, H. C., T. C. Pagano, S. Sorooshian, and R. Bales, 2002: Confidence

References

- builders. Evaluating seasonal climate forecasts from user perspectives. *Bull. Amer. Meteorol. Soc.*, **83**, 683–698.
- Hastenrath, S., 1984: Predictability of north-east Brazil droughts. *Nature*, **307**, 531–533.
- Hisdal, H., K. Stahl, L. M. Tallaksen, and S. Demuth, 2001: Have streamflow droughts in Europe become more severe or frequent? *Int. J. Climat.*, **21**, 317–333.
- Hoerling, M. P., J. W. Hurrell, and T. Xu, 2001: Tropical origins for recent North Atlantic climate change. *Science*, **292**, 90–92.
- Houghton, J., Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, and D. Xiaosu, 2001: *Climate Change 2001: The Scientific Basis*. Cambridge University Press, U.K., 944. IPCC Third Assessment Report.
- Hsieh, W. W. and B. Tang, 1998: Applying neural network models to prediction and data analysis in meteorology and oceanography. *Bull. Amer. Meteorol. Soc.*, **79**, 1855–1870.
- Hurrell, J. W., 1995: Decadal trends in the North Atlantic Oscillation: regional temperature and precipitation. *Science*, **269**, 676–679.
- Hurrell, J. W., M. P. Hoerling, A. S. Phillips, and T. Xu, 2004: Twentieth century North Atlantic climate change. Part I: assessing determinism. *Climate Dyn.*, **23**, 371–389.
- Hurrell, J. W., Y. Kushnir, G. Ottersen, and M. Visbeck, 2003: An overview of the North Atlantic Oscillation. In *North Atlantic Oscillation: Climate significance and environmental impact*, Hurrell, J. W., Kushnir, Y., Ottersen, G., and Visbeck, M., editors, volume 134. American Geophysical Union. Geophysical Monograph Series, 1–35.
- Hurrell, J. W. and H. van Loon, 1997: Decadal variations in climate associated with the North Atlantic Oscillation. *Climatic Change*, **36**, 301–326.
- Johansson, A., A. Barnston, S. Saha, and H. van den Dool, 1998: On the level and origin of seasonal forecast skill in northern Europe. *J. Atmos. Sci.*, **55**, 103–127.
- Jones, P. D., 1994: Hemispheric surface air temperature variations: A reanalysis and an update to 1993. *J. Climate*, **7**, 1794–1802.

- Jones, P. D., T. Jónsson, and D. Wheeler, 1997: Extension to the North Atlantic Oscillation using early instrumental pressure observations from Gibraltar and southwest Iceland. *Int. J. Climat.*, **17**, 1433–1450.
- Jones, P. D. and A. Moberg, 2003: Hemispheric and large-scale surface air temperature variations: An extensive revision and an update to 2001. *J. Climate*, **16**, 206–223.
- Jones, P. D., M. New, D. E. Parker, S. Martin, and I. G. Rigor, 1999: Surface air temperature and its changes over the last 150 years. *Rev. Geophys.*, **37**, 173–199.
- Jones, P. D., T. J. Osborn, K. R. Briffa, C. K. Folland, E. B. Horton, L. V. Alexander, D. E. Parker, and N. A. Rayner, 2001: Adjusting for sampling density in grid-box land and ocean surface temperature time series. *J. Geophys. Res.*, **106**, 3371–3380.
- Jung, T. and M. Hilmer, 2001: The link between the North Atlantic Oscillation and Arctic sea ice export through Fram Strait. *J. Climate*, **14**, 3932–3943.
- Junge, M. M. and D. B. Stephenson, 2003: Mediated and direct effects of the North Atlantic ocean on winter temperatures in northwest Europe. *Int. J. Climat.*, **23**, 245–261.
- Kalnay, E., M. Kanamitsu, and coauthors, 1996: The NCEP/NCAR 40-year Reanalysis Project. *Bull. Amer. Meteorol. Soc.*, **77**, 437–471.
- Kanamitsu, M., A. Kumar, and Coauthors, 2002: NCEP dynamical seasonal forecast system 2000. *Bull. Amer. Meteorol. Soc.*, **83**, 1019–1037.
- Kaplan, A., M. A. Cane, Y. Kushnir, A. C. Clement, M. B. Blumenthal, and B. Rajagopalan, 1998: Analyses of global sea surface temperature 1856–1991. *J. Geophys. Res.*, **C9**, 18,567–18,589.
- Kaplan, A., Y. Kushnir, and M. A. Cane, 2000: Reduced space optimal interpolation of historical marine sea level pressure: 1854–1992. *J. Climate*, **13**, 2987–3002.
- Kodera, K. and M. Chiba, 1989: Western Siberian spring snow cover and East Asian June 500 mb height. *Papers Meteorol. Geophys.*, **40**, 51–54.
- Krishnamurti, T. N., C. M. Kishtawal, T. E. LaRow, D. R. Bachiochi, Z. Zhang, C. E. Williford, S. Gadgil, and S. Surendran, 1999: Improved weather and seasonal climate forecasts from multimodel superensemble. *Science*, **285**, 1548–

- 1550.
- Lau, K. and H. Weng, 1995: Climate signal detection using wavelet transformation: How to make a time series sing. *Bull. Amer. Meteorol. Soc.*, **76**, 2391–2402.
- Levinson, D. H. and A. M. Waple, 2004: State of climate in 2003. *Bull. Amer. Meteorol. Soc.*, **85**, S1–S72.
- Lin, Z., H. Weng, G. Zhou, H. Chen, X. Lang, Y. Zhao, and Q. Zeng, 2004: Recent advances in dynamical extra-seasonal to annual climate prediction at IAP/CAS. *Adv. in Atmos. Sci.*, **21**, 456–466.
- Liston, G. E., 2004: Representing subgrid snow cover heterogeneities in regional and global models. *J. Climate*, **17**, 1381–1397.
- Liu, H., Z. Sun, J. Wang, and J. Min, 2004: A modelling study of the effects of anomalous snow cover on the Tibetan Plateau upon the South Asian summer monsoon. *Adv. in Atmos. Sci.*, **21**, 964–975.
- Livezey, R. E. and W. Y. Chen, 1983: Statistical field significance and its determination by Monte Carlo techniques. *Mon. Wea. Rev.*, **111**, 46–59.
- Lloyd-Hughes, B., 2002: *The Long-Range Predictability of European Drought*. PhD thesis, University College London.
- Lloyd-Hughes, B., M. A. Saunders, and P. Rockett, 2004: A consolidated CLIPER model for improved August–September ENSO prediction skill. *Weather and Forecasting*, **19**, 1089–1105.
- Lorenz, E. N., 1969: The predictability of flow which possesses many scales of motion. *Tellus*, **21**, 289–307.
- Lorenz, E. N., 1970: Climate change as a mathematical problem. *J. Appl. Meteor.*, **9**, 325–329.
- Luterbacher, J., E. Xoplaki, and Coauthors, 2002: Extending NAO reconstructions back to 1500. *Atmos. Sci. Lett.*, **2**, 114–124.
- Manley, G., 1974: Central England Temperatures: monthly means 1659 to 1973. *Quart. J. Roy. Meteor. Soc.*, **100**, 389–405.
- Manly, B. F. J., 1997: *Randomization, Bootstrap and Monte Carlo Methods in Biology*. Chapman and Hall, London. 2nd ed, 424pp.
- Marshall, J. and Y. K. Coauthors, 2001: North Atlantic climate variability: phe-

- nomena, impacts and mechanisms. *Int. J. Climat.*, **21**, 1863–1898.
- Mathieu, P.-P., R. T. Sutton, B. Dong, and M. Collins, 2004: Predictability of winter climate over the North Atlantic European region during ENSO affairs. *J. Climate*, **17**, 1953–1973.
- Mercer, J. B., 2003: Cold - an underrated risk factor for health. *Environmental Research*, **92**, 8–13.
- Michaelsen, J., 1987: Cross-validation in statistical climate forecast models. *J. Climate Appl. Meteor.*, **26**, 1589–1600.
- Mimmack, G. M., S. J. Mason, and J. S. Galpin, 2001: Distance metrics in hierarchical cluster analysis: defining regions. *J. Climate*, **14**, 2790–2797.
- Morrissey, M. L., 1990: An evaluation of ship data in the equatorial Pacific. *J. Climate*, **3**, 99–112.
- Oberhuber, J., 1993: Simluation of the Atlantic circulation with a coupled sea ice - mixed layer - isopycnal general circulation model. Part I: Model description. *J. Phys. Oceanogr.*, **23**, 808–829.
- Osborn, T. J., K. R. Briffa, S. F. B. Tett, P. D. Jones, and R. M. Trigo, 1999: Evaluation of the North Atlantic Oscillation as simulated by a coupled climate model. *Climate Dyn.*, **15**, 685–702.
- Ose, T., 1996: The comparison of the simulated response to the regional snowmass anomalies over Tibet, Eastern Europe and Siberia. *J. Meteor. Soc. Japan*, **74**, 845–866.
- Paeth, H., A. Hense, R. Glowienka-Hense, R. Voss, and U. Cubasch, 1999: The North Atlantic Oscillation as an indicator for greenhouse-gas induced regional climate change. *Climate Dyn.*, **15**, 953–960.
- Paeth, H., M. Latif, and A. Hense, 2003: Global SST influence on twentieth century NAO variability. *Climate Dyn.*, **21**, 63–75.
- Palmer, T. N., A. Alessandri, and coauthors, 2004: Development of a European multimodel ensemble system for seasonal-to-interannual prediction (DEMETER). *Bull. Amer. Meteorol. Soc.*, **85**, 853–872.
- Palmer, T. N. and D. L. T. Anderson, 1994: The prospects for seasonal forecasting - a review paper. *Quart. J. Roy. Meteor. Soc.*, **120**, 755–793.
-

References

- Palmer, T. N., C. Brankovic, and D. S. Richardson, 2000: A probability and decision-model analysis of PROVOST seasonal multi-model ensemble integrations. *Quart. J. Roy. Meteor. Soc.*, **126**, 2013–2034.
- Parker, D. E., C. K. Folland, and M. Jackson, 1995: Marine surface temperature: observed variations and data requirements. *Climatic Change*, **31**, 559–600.
- Peng, S., W. A. Robinson, and S. Li, 2003: Mechanisms for the NAO response to the North Atlantic tripole. *J. Climate*, **16**, 1987–2004.
- Peng, S. and J. S. Whittaker, 1999: Mechanisms determining the atmospheric response to midlatitude SST anomalies. *J. Climate*, **12**, 1393–1408.
- Pielke, R., G. E. Liston, W. L. Chapman, and D. A. Robinson, 2004: Actual and insolation-weighted Northern Hemisphere snow cover and sea ice between 1973–2002. *Climate Dyn.*, **22**, 591–595.
- Qian, B. and M. A. Saunders, 2003: Summer U.K. temperature and its links to preceding Eurasian snow cover, North Atlantic SSTs, and the NAO. *J. Climate*, **16**, 4108–4120.
- Radcliffe, R. A. S. and R. Murray, 1970: New lag associations between North Atlantic sea temperature and European pressure applied to long-range weather forecasting. *Quart. J. Roy. Meteor. Soc.*, **96**, 226–246.
- Randall, D. A., R. D. Hess, and Coauthors, 1994: Analysis of snow feedbacks in fourteen general circulation models. *J. Geophys. Res.*, **99**, 20757–20771.
- Rayner, N. A., E. B. Horton, D. E. Parker, C. K. Folland, and R. B. Hackett, 1996: *Version 2.2 of the Global sea-Ice and Sea Surface Temperature data set, 1903–1994*. CRTN 74, Hadley Centre for Climate Prediction and Research, Meteorological Office, London Road, Bracknell, Berkshire, RG12 2SY. 21 pp plus figures.
- Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. K. Alexander, and D. P. Rowell, 2003: Global analyses of sea surface temperature, sea ice and night marine air temperature since the late nineteenth century. *J. Geophys. Res.*, **108**, 4407, doi:10.1029/2002JD002670.
- Reynolds, R. W., 1988: A real-time global sea surface temperature analysis. *J. Climate*, **1**, 75–86.

- Reynolds, R. W., N. Rayner, T. Smith, D. Stokes, and W. Wang, 2002: An improved in situ and satellite SST analysis. *J. Climate*, **15**, 1609–1625.
- Reynolds, R. W. and T. M. Smith, 1994: Improved global sea surface temperature analyses using optimum interpolation. *J. Climate*, **7**, 929–948.
- Reynolds, R. W. and T. M. Smith, 1995: A high resolution global sea surface temperature climatology. *J. Climate*, **8**, 1571–1583.
- Robertson, A. W., 2000: The influence of Atlantic sea surface temperature anomalies on the North Atlantic Oscillation. *J. Climate*, **13**, 122–138.
- Robinson, D. A., K. F. Dewey, and R. R. Heim, 1993: Global snow cover monitoring: An update. *Bull. Amer. Meteorol. Soc.*, **74**, 1689–1696.
- Robock, A., M. Mingquan, K. Y. Vinnikov, and D. Robinson, 2003: Land surface conditions over Eurasia and Indian summer monsoon rainfall. *J. Geophys. Res.*, **108(D4)**, doi:10.1029/2002JD002286.
- Rodwell, M. J. and C. K. Folland, 2002: Atlantic air-sea interaction and seasonal predictability. *Quart. J. Roy. Meteor. Soc.*, **128**, 1413–1443.
- Rodwell, M. J., D. P. Rowell, and C. K. Folland, 1999: Oceanic forcing of the wintertime North Atlantic Oscillation and European climate. *Nature*, **398**, 320–323.
- Roeckner, E., L. Bengtsson, J. Feichter, J. Lelieveld, and H. Rodhe, 1999: Transient climate change simulations with a coupled atmosphere-ocean GCM including the tropospheric sulfur cycle. *J. Climate*, **12**, 3004–3032.
- Roeckner, E., J. M. Oberhuber, A. Bacher, M. Christoph, and I. Kirchner, 1996: ENSO variability and atmospheric response in a global coupled atmosphere-ocean GCM. *Climate Dyn.*, **12**, 737–754.
- Rogers, J. C. and H. van Loon, 1979: The seesaw in winter temperatures between Greenland and northern Europe. Part II: Some oceanic and atmospheric effects in middle and high latitudes. *Mon. Wea. Rev.*, **107**, 509–519.
- Sahai, A. K., M. K. Soman, and V. Satyan, 2000: All India summer monsoon rainfall prediction using an artificial neural network. *Climate Dynamics*, **16**, 291–302.
- Saito, K., J. Cohen, and D. Entekhabi, 2001: Evolution of atmospheric response to

- early-season snow cover anomalies. *Mon. Wea. Rev.*, **129**, 2746–2760.
- Saito, K., T. Yasunari, and J. Cohen, 2004: Changes in the sub-decadal covariability between Northern Hemisphere snow cover and the general circulation of the atmosphere. *Int. J. Climat.*, **24**, 33–44.
- Saunders, M. A. and A. R. Harris, 1997: Statistical evidence links exceptional 1995 Atlantic hurricane season to record sea warming. *Geophysical Research Letters*, **10**, 1255–1258.
- Saunders, M. A. and B. Qian, 2002: Seasonal predictability of the winter NAO from North Atlantic sea surface temperatures. *Geophys. Res. Lett.*, **29**, doi:10.1029/2002GL014952.
- Saunders, M. A., B. Qian, and B. Lloyd-Hughes, 2003: Summer snow extent heralding of the winter North Atlantic Oscillation. *Geophys. Res. Lett.*, **30**, doi:10.1029/2003GL017401.
- Saur, J. F. T., 1963: A study of the quality of sea water temperatures reported in logs of ships' weather observations. *Journal of Applied Meteorology*, **2**, 417–427.
- Shinoda, M., 2001: Climate memory of snow mass as soil moisture over central Eurasia. *J. Geophys. Res.*, **106**, 33,393–33,403.
- Shukla, J., 1998: Predictability in the midst of chaos: A scientific basis for climate forecasting. *Science*, **282**, 728–731.
- Skaugen, T. E. and O. E. Tveito, 2004: Growing-season and degree-day scenario in Norway for 2021–2050. *Climate Res.*, **26**, 221–232.
- Slater, A. G., C. A. Schlosser, and Coauthors, 2001: The representation of snow in land surface schemes: Results from PILPS 2(d). *J. Hydrometeor.*, **2**, 7–25.
- Smith, T. M. and R. W. Reynolds, 2004a: Improved extended reconstruction of SST (1854–1997). *J. Climate*, **17**, 2466–2477.
- Smith, T. M. and R. W. Reynolds, 2004b: Reconstruction of monthly mean oceanic sea level pressure based on COADS and station data (1854–1997). *J. Atmos. Ocean. Tech.*, **21**, 1272–1282.
- Spencer, R. W. and J. R. Christy, 1990: Precise monitoring of global temperature trends from satellites. *Science*, **247**, 1558–1662.
- Sutton, R. T. and D. L. R. Hodson, 2003: Influence of the ocean on North Atlantic

- climate variability 1871-1999. *J. Climate*, **16**, 3296-3313.
- Sutton, R. T., W. A. Norton, and S. P. Jewson, 2001: The North Atlantic Oscillation - what role for the ocean? *Atmos. Sci. Lett.*, **1**, 89-100.
- Thompson, D. W. J. and J. M. Wallace, 1998: The Arctic Oscillation signature in the wintertime geopotential height and temperature fields. *Geophys. Res. Lett.*, **25**, 1297-1300.
- Thompson, D. W. J. and J. M. Wallace, 2000: Annular modes in the extratropical circulation. Part I: Month-to-month variability. *J. Climate*, **13**, 1000-1016.
- Torrence, C. and G. P. Compo, 1998: A practical guide to wavelet analysis. *Bull. Amer. Meteorol. Soc.*, **79**, 61-78.
- Trenberth, K. E., G. W. Branstator, D. Karoly, A. Kumar, N. C. Lau, and C. Ropelewski, 1998: Progress during TOGA in understanding and modeling global teleconnections associated with tropical sea surface temperatures. *J. Geophys. Res.*, **103**, 14291-14324.
- Trenberth, K. E., J. R. Christy, and J. W. Hurrell, 1992: Monitoring global monthly mean surface temperatures. *J. Climate*, **5**, 1405-1423.
- Trigo, R. M., T. J. Osborn, and J. M. Corte-Real, 2002: The North Atlantic Oscillation influence on Europe: climate impacts and associated physical mechanisms. *Clim. Res.*, **20**, 9-17.
- Trigo, R. M., D. Poso-Vázquez, T. J. Osborn, Y. Castro-Díez, S. Gámiz-Fortis, and M. J. Esteban-Parra, 2004: North Atlantic Oscillation influence on precipitation, river flow and water resources in the Iberian Peninsula. *Int. J. Climat.*, **24**, 925-944.
- UKMO, 2004: *Global Seasonal Forecasts*. United Kingdom Meteorological Office. <http://www.metoffice.com/research/seasonal/index.html/>.
- van den Dool, H. M., 1983: A possible explanation of the observed persistence of monthly mean circulation anomalies. *Mon. Wea. Rev.*, **111**, 539-544.
- van den Dool, H. M. and J. L. Nap, 1981: An explanation of persistence in monthly mean temperatures in the Netherlands. *Tellus*, **33**, 123-131.
- van den Dool, H. M. and J. L. Nap, 1985: Short and long range air temperature forecasts near an ocean. *Mon. Wea. Rev.*, **113**, 878-886.

- van den Dool, H. M. and S. Saha, 1990: Frequency dependence in forecast skill. *Mon. Wea. Rev.*, **118**, 128–137.
- van Oldenborgh, G. J., M. A. Balmaseda, L. Ferranti, T. N. Stockdale, and D. L. T. Anderson, 2003: Did the ECMWF seasonal forecast model outperform a statistical model over the last 15 years? Technical Memoranda 404, ECMWF, Shinfield Park, Reading, U.K.
- Vernekar, A. D., J. Zhou, and J. Shukla, 1995: The effects of Eurasian snow cover on the Indian monsoon. *J. Climate*, **8**, 248–266.
- von Storch, H. and F. W. Zwiers, 1999: *Statistical Analysis in Climate Research*. Cambridge University Press, Cambridge.
- Walker, G. T., 1924: Correlation of seasonal variations in weather IX: A further study of world weather. *Members of the Indian Meteorological Department*, **24**, 275–332.
- Walker, G. T. and E. W. Bliss, 1932: World weather V. *Mem. Roy. Met. Soc.*, **4**, 53–84.
- Wallace, J. M., 2000: North Atlantic Oscillation/annular mode: Two paradigms – one phenomenon. *Quart. J. Roy. Meteor. Soc.*, **126**, 791–805.
- Wallace, J. M., E. N. Rasmusson, T. P. Mitchell, V. E. Kousky, E. S. Sarachik, and H. von Storch, 1998: The structure and evolution of ENSO-related climate variability in the tropical Pacific: Lessons from TOGA. *J. Geophys. Res.*, **103**, 14241–14259.
- Walsh, J. E., W. H. Jasperson, and B. Ross, 1985: Influences of snow cover and soil moisture on monthly air temperature. *Mon. Wea. Rev.*, **113**, 756–768.
- Wanner, H., S. Brönnimann, C. Casty, D. Gyalistras, J. Luterbacher, C. Schmoltz, D. B. Stephenson, and E. Xoplaki, 2001: North Atlantic Oscillation - concepts and studies. *Surveys in Geophysics*, **22**, 321–382.
- Watanabe, M. and T. Nitta, 1998: Relative impact of snow and sea surface temperature anomalies on an extreme phase in the winter atmospheric circulation. *J. Climate*, **11**, 2837–2857.
- Watanabe, M. and T. Nitta, 1999: Decadal changes in the atmospheric circulation and associated surface climate variations in the Northern Hemisphere winter. *J.*

- Climate*, **12**, 494–510.
- Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences*. Academic Press, London. 467pp.
- WMO, 2002: *Standardised Verification System (SVS) for Long-Range Forecasts (LRF)*. New attachment II-9 to the Manual on the GDPS (WMO-No.485), Volume 1, WMO, Geneva, 21pp.
- Woodruff, S. D., H. Diaz, J. Elms, and S. Worley, 1998: COADS Release 2 data and metadata enhancements for improvements of marine surface flux fields. *Phys. Chem. Earth*, **23**, 517–527.
- Zwiers, F. W. and H. von Storch, 2004: On the role of statistics in climate research. *Int. J. Climat.*, **24**, 665–680.